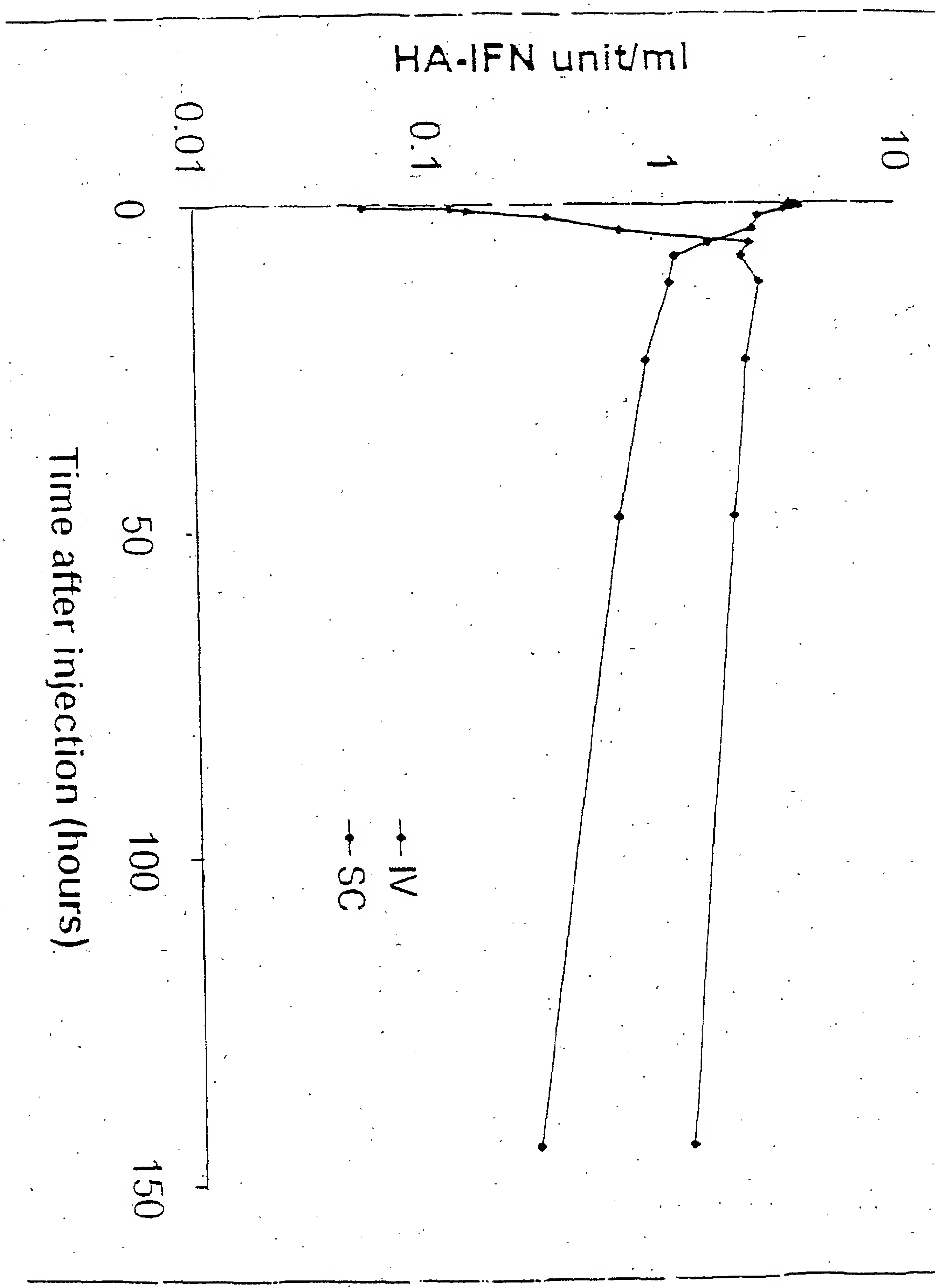

HA-IFN unit/ml
0.01
0.1
1
10
0
50
100
150
Time after injection (hours)
IV
SC

Figure 6

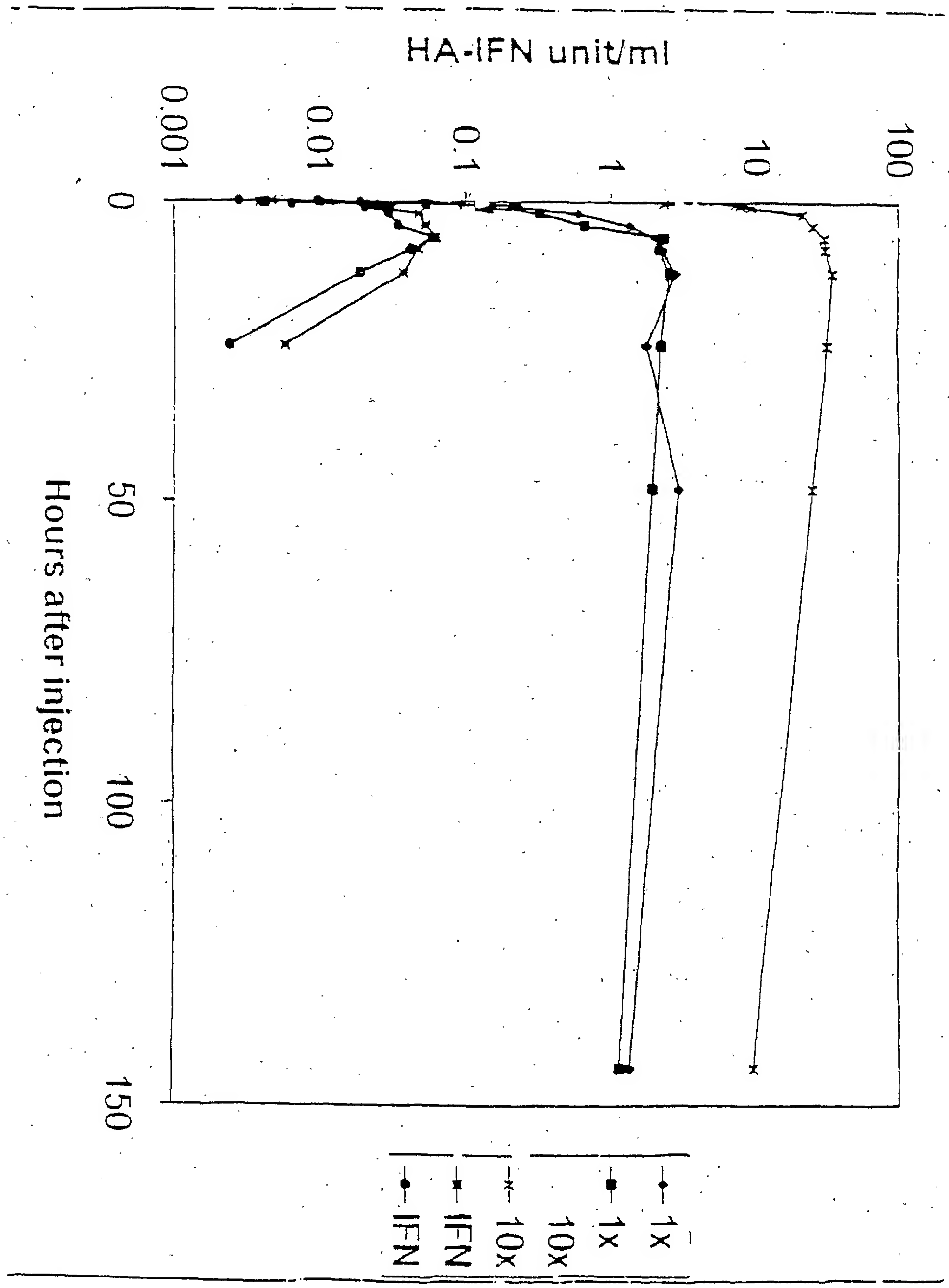

HA-IFN unit/ml
100
10
1
0.1
0.01
0.001
0
50
100
150
Hours after injection
1x
1x
10x
10x
IFN
IFN

Figure 7

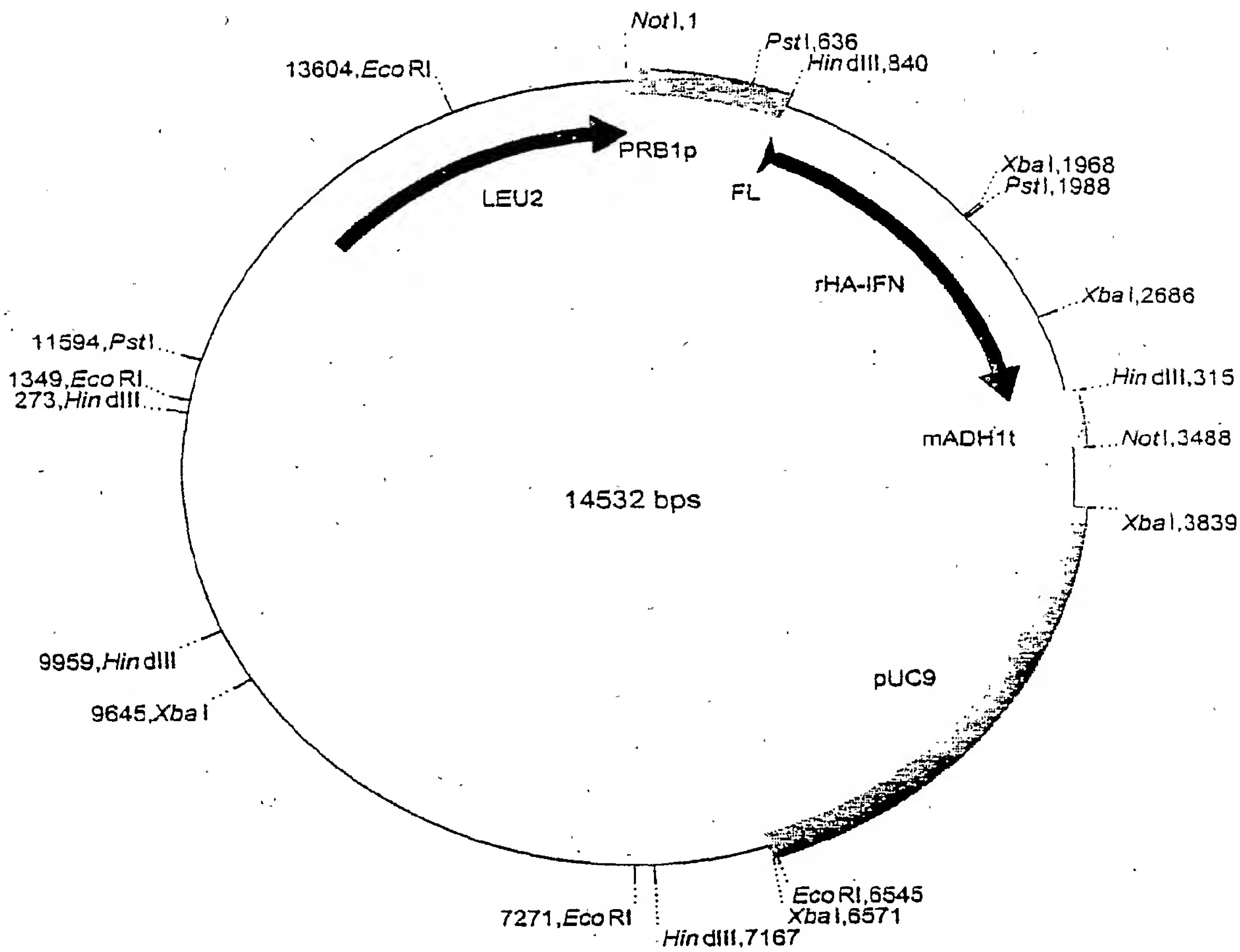


Figure 8. The HA-IFNα expression cassette in pSAC35. The expression cassette comprises
*PRB*1 promoter, from *S. cerevisiae*.
Fusion leader, first 19 amino acids of the HA leader followed by the last 6 amino acids of the MFα-1 leader.
HA-IFNα coding sequence with a double stop codon (TAATAA)
*ADH*1 terminator, from *S. cerevisiae*. Modified to remove all the coding sequence normaly present in the *Hin*d III/*Bam*HI fragment generally used.

Figure 8

Localisation of 'Loops' based on the HA Crystal Structure which could be used for Mutation/Insertion

```
                I                               II              III
1      DAHKSEVAHR FKDLGEENFK ALVLIAFAQY LQQCPFEDHV KLVNEVTEFA
            HHHHH HHH    HHH HHHHHHHHHH      HHHHH HHHHHHHHHH

                I                          II              III
51     KTCVADESAE NCDKSLHTLF GDKLCTVATL RETYGEMADC CAKQEPERNE
       HHHHH           HHHHH HHHHH            HHHH H     HHHH

101    CFLQHKDDNP NLPRLVRPEV DVMCTAFHDN EETFLKKYLY EIARRHPYFY
       HHHH                H HHHHHHHH    HHHHHHHHH HHHHH

                                IV
151    APELLFFAKR YKAAFTECCQ AADKAACLLP KLDELRDEGK ASSAKQRLKC
       HHHHHHHHHH HHHHHHHHH       HHHHH HHHEHHHHHH HHHHHHHHHH

                                                          V
201    ASLQKFGERA FKAWAVARLS QRFPKAEFAE VSKLVTDLTK VHTECCHGDL
       HHHHH   HH HHHHHHHHHH HH     HHH HHHHHHHHHH HHHHHH  HH

                            VI              VII
251    LECADDRADL AKYICENQDS ISSKLKECCE KPLLEKSHCI AEVENDEMPA
       HHHHHHHHHH HHHHH          HHHHH     HHHHHHH H

301    DLPSLAADFV ESKDVCKNYA EAKDVFLGMF LYEYARRHPD YSVVLLLRLA
            HHHH      HHHHHH    HHHHHHH HHHHHH       HHHHHHHH

                   VIII
351    KTYETTLEKC CAAADPHECY AKVFDEFKPL VEEPQNLIKQ NCELFEQLGE
       HHHHHHHHHH       HH    H    HHHHH HHHHHHHHHH HHHHHHH

                                                    IX
401    YKFQNALLVR YTKKVPQVST PTLVEVSRNL GKVGSKCCKH PEAKRMPCAE
       HHHHHHHHHH HHHH     H HHHHHHHHHH    HHH       HHHHHHHH

                        X                    XI
451    DYLSVVLNQL CVLHEKTPVS DRVTKCCTES LVNRRPPCFSA LEVDETYVPK
       HHHHHHHHHH HHHHH      HHHHHHHHH     HHHHHHHH

501    EFNAETFTFH ADICTLSEKE RQIKKQTALV ELVKHKPKAT KEQLKAVMDD
                   HHH   HHH HHHHMMEHHH HHH          HHHHHHHH

                    XII
551    FAAFVEKCCK ADDKETCFAE EGKKLVAASQ AALGL
       HHHHHHHH         HHHH HHHHHHHHHH HH
```

Loop		Loop	
I	Val54-Asn61	VII	Glu280-His288
II	Thr76-Asp89	VIII	Ala362-Glu368
III	Ala92-Glu100	IX	Lys439-Pro447
IV	Gln170-Ala176	X	Val462-Lys475
V	His247-Glu252	XI	Thr478-Pro486
VI	Glu266-Glu277	XII	Lys560-Thr566

Examples of Modifications to Loop IV

a. Randomisation of Loop IV.

```
                                   IV
151   APELLFFAKR YKAAFTECCQ AADKAACLLP KLDELRDEGK ASSAKQRLKC
      HHHHHHHHHH HHHHHHHHH        HHHHH HHHHHHHHHH HHHHHHHHHH

                                   IV
151   APELLFFAKR YKAAFTECCX XXXXXXCLLP KLDELRDEGK ASSAKQRLKC
      HHHHHHHHHH HHHHHHHHH        HHHHH HHHHHHHHHH HHHHHHHHHH
```

X represents the mutation of the natural amino acid to any other amino acid. One, more or all of the amino acids can be changed in this manner. This figure indicates all the residues have been changed.

b. Insertion (or replacement) of Randomised sequence into Loop IV.

```
                                      (X)n
                                       ↓
                                      IV
151   APELLFFAKR YKAAFTECCQ AADKAACLLP KLDELRDEGK ASSAKQRLKC
      HHHHHHHHHH HHHHHHHHH        HHHHH HHHHHHHHHH HHHHHHHHHH
```

The insertion can be at any point on the loop and the length a length where n would typically be 6, 8, 12, 20 or 25.

Figure 10

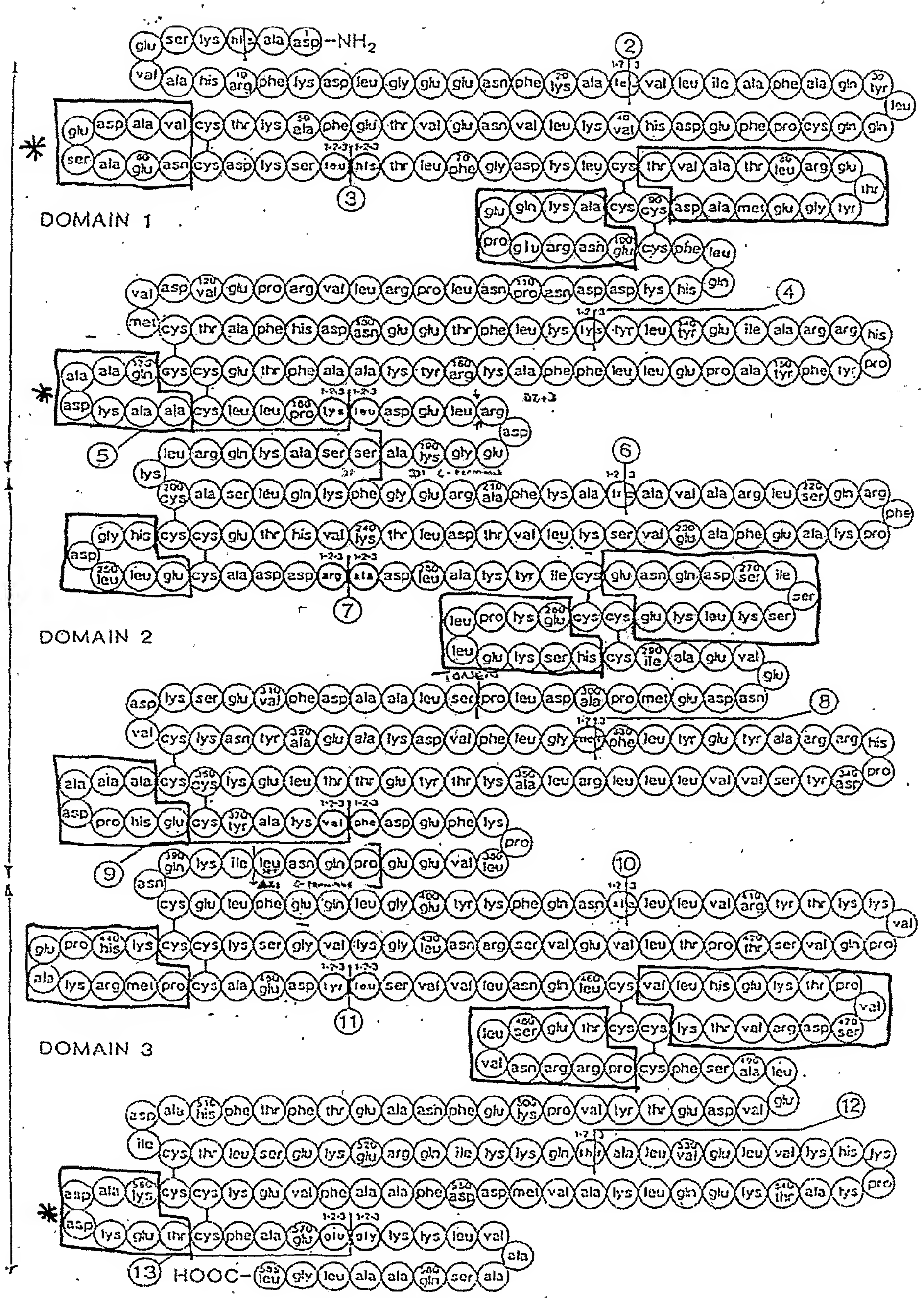
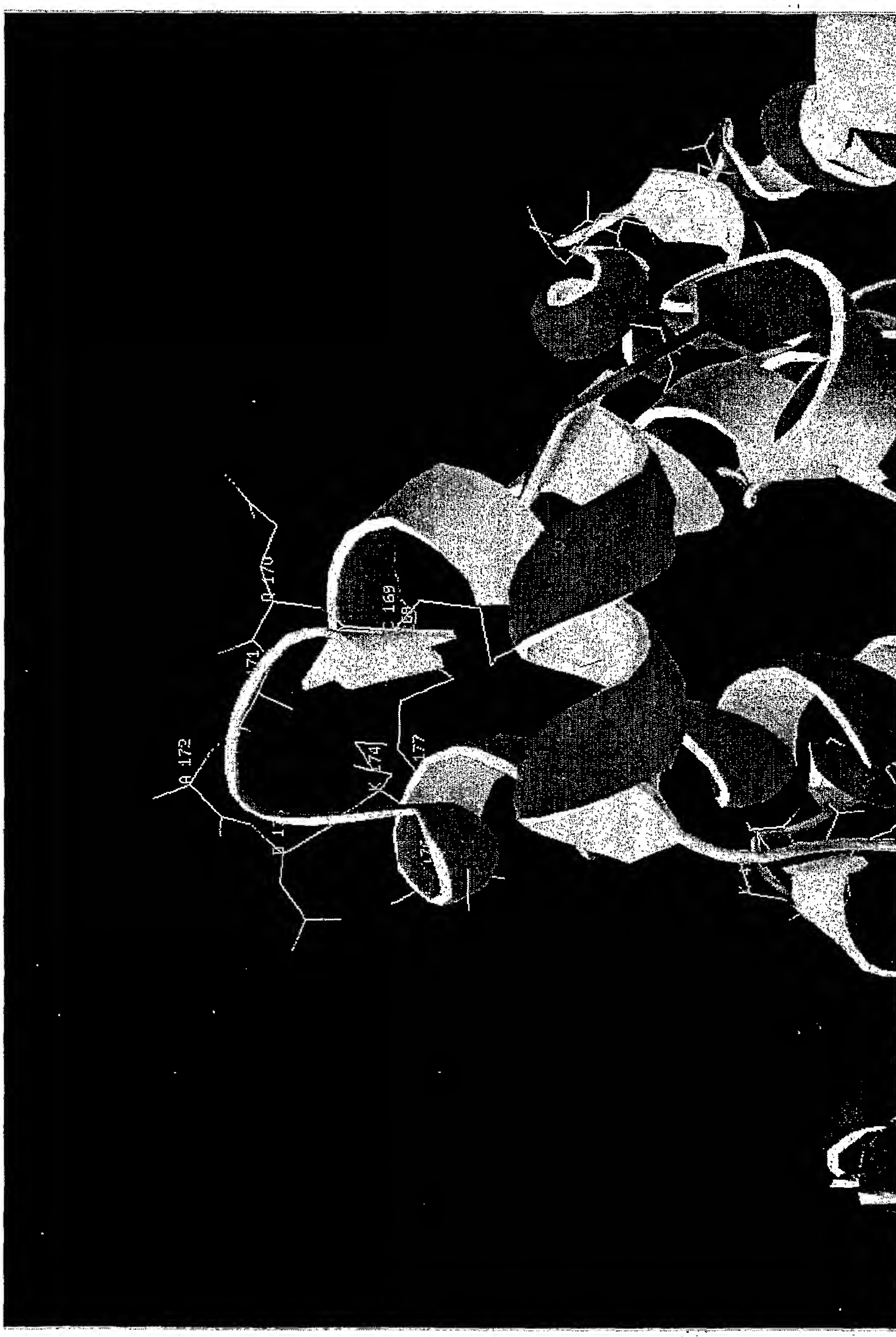


Figure 11



Disulfide bonds shown in yellow

Figure 12: Loop IV Gln170-Ala176

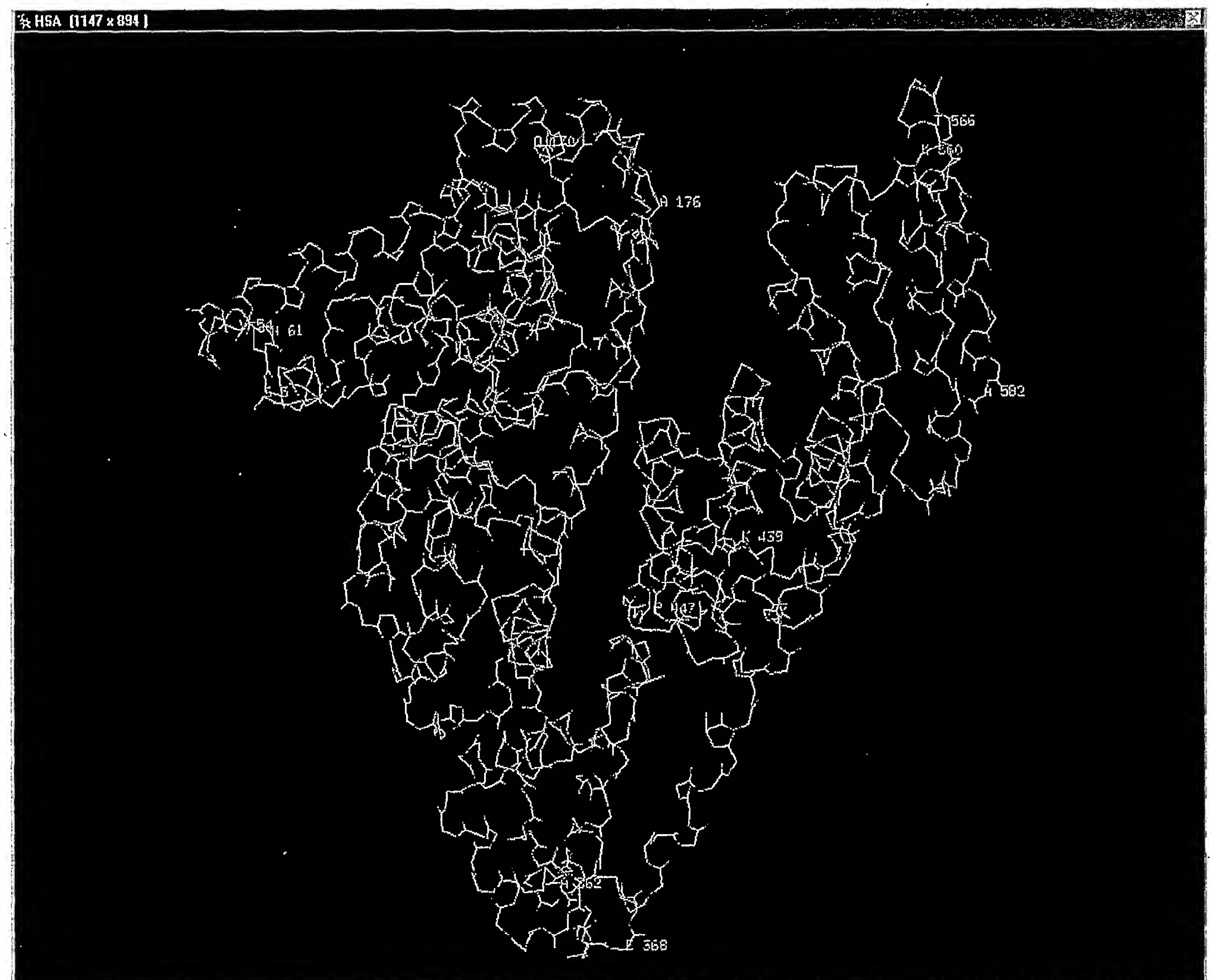


Figure 13: Tertiary Structure of HA

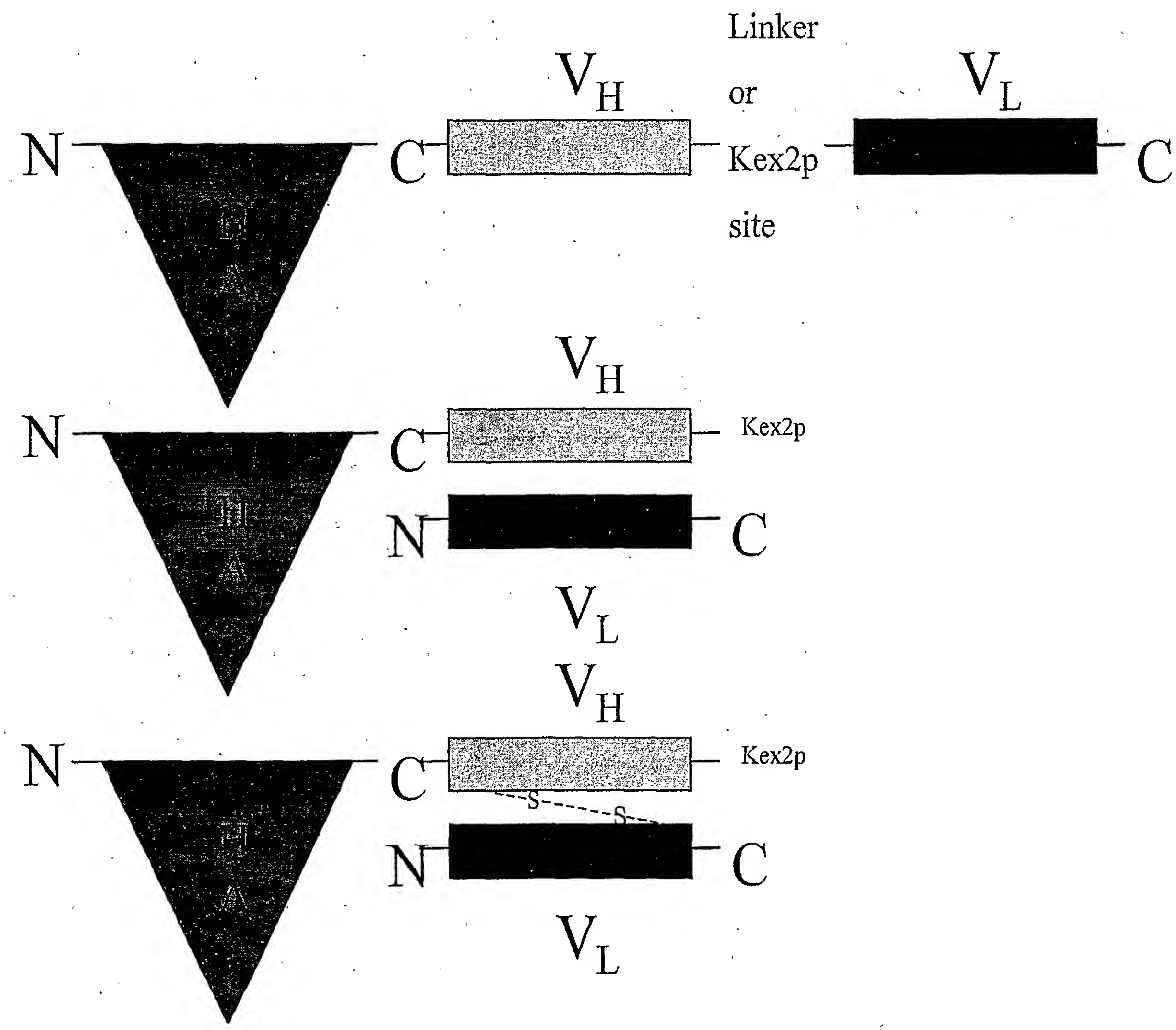


Figure 14: Schematic Diagram of Possible ScFv Fusions (Example is of a C-terminal fusion to HA)

```
  1 GAT GCA CAC AAG AGT GAG GTT GCT CAT CGG TTT AAA GAT TTG GGA GAA GAA AAT TTC AAA 60
  1 D   A   H   K   S   E   V   A   H   R   F   K   D   L   G   E   E   N   F   K   20

 61 GCC TTG GTG TTG ATT GCC TTT GCT CAG TAT CTT CAG CAG TGT CCA TTT GAA GAT CAT GTA 120
 21 A   L   V   L   I   A   F   A   Q   Y   L   Q   Q   C   P   F   E   D   H   V   40

121 AAA TTA GTG AAT GAA GTA ACT GAA TTT GCA AAA ACA TGT GTT GCT GAT GAG TCA GCT GAA 180
 41 K   L   V   N   E   V   T   E   F   A   K   T   C   V   A   D   E   S   A   E   60

181 AAT TGT GAC AAA TCA CTT CAT ACC CTT TTT GGA GAC AAA TTA TGC ACA GTT GCA ACT CTT 240
 61 N   C   D   K   S   L   H   T   L   F   G   D   K   L   C   T   V   A   T   L   80

241 CGT GAA ACC TAT GGT GAA ATG GCT GAC TGC TGT GCA AAA CAA GAA CCT GAG AGA AAT GAA 300
 81 R   E   T   Y   G   E   M   A   D   C   C   A   K   Q   E   P   E   R   N   E   100

301 TGC TTC TTG CAA CAC AAA GAT GAC AAC CCA AAC CTC CCC CGA TTG GTG AGA CCA GAG GTT 360
101 C   F   L   Q   H   K   D   D   N   P   N   L   P   R   L   V   R   P   E   V   120

361 GAT GTG ATG TGC ACT GCT TTT CAT GAC AAT GAA GAG ACA TTT TTG AAA AAA TAC TTA TAT 420
121 D   V   M   C   T   A   F   H   D   N   E   E   T   F   L   K   K   Y   L   Y   140

421 GAA ATT GCC AGA AGA CAT CCT TAC TTT TAT GCC CCG GAA CTC CTT TTC TTT GCT AAA AGG 480
141 E   I   A   R   R   H   P   Y   F   Y   A   P   E   L   L   F   F   A   K   R   160
```

Figure 15A

```
481 TAT AAA GCT GCT TTT ACA GAA TGT TGC CAA GCT GCT GAT AAA GCT GCC TGC CTG TTG CCA 540
161 Y   K   A   A   F   T   E   C   C   Q   A   A   D   K   A   A   C   L   L   P   180

541 AAG CTC GAT GAA CTT CGG GAT GAA GGG AAG GCT TCG TCT GCC AAA CAG AGA CTC AAA TGT 600
181 K   L   D   E   L   R   D   E   G   K   A   S   S   A   K   Q   R   L   K   C   200

601 GCC AGT CTC CAA AAA TTT GGA GAA AGA GCT TTC AAA GCA TGG GCA GTG GCT CGC CTG AGC 660
201 A   S   L   Q   K   F   G   E   R   A   F   K   A   W   A   V   A   R   L   S   220

661 CAG AGA TTT CCC AAA GCT GAG TTT GCA GAA GTT TCC AAG TTA GTG ACA GAT CTT ACC AAA 720
221 Q   R   F   P   K   A   E   F   A   E   V   S   K   L   V   T   D   L   T   K   240

721 GTC CAC ACG GAA TGC TGC CAT GGA GAT CTG CTT GAA TGT GCT GAT GAC AGG GCG GAC CTT 780
241 V   H   T   E   C   C   H   G   D   L   L   E   C   A   D   D   R   A   D   L   260

781 GCC AAG TAT ATC TGT GAA AAT CAG GAT TCG ATC TCC AGT AAA CTG AAG GAA TGC TGT GAA 840
261 A   K   Y   I   C   E   N   Q   D   S   I   S   S   K   L   K   E   C   C   E   280

841 AAA CCT CTG TTG GAA AAA TCC CAC TGC ATT GCC GAA GTG GAA AAT GAT GAG ATG CCT GCT 900
281 K   P   L   L   E   K   S   H   C   I   A   E   V   E   N   D   E   M   P   A   300

901 GAC TTG CCT TCA TTA GCT GCT GAT TTT GTT GAA AGT AAG GAT GTT TGC AAA AAC TAT GCT 960
301 D   L   P   S   L   A   A   D   F   V   E   S   K   D   V   C   K   N   Y   A   320
```

Figure 15B

```
961 GAG GCA AAG GAT GTC TTC CTG GGC ATG TTT TTG TAT GAA TAT GCA AGA AGG CAT CCT GAT 1020
321 E   A   K   D   V   F   L   G   M   F   L   Y   E   Y   A   R   R   H   P   D   340

1021 TAC TCT GTC GTG CTG CTG CTG AGA CTT GCC AAG ACA TAT GAA ACC ACT CTA GAG AAG TGC 1080
341 Y   S   V   V   L   L   L   R   L   A   K   T   Y   E   T   T   L   E   K   C   360

1081 TGT GCC GCT GCA GAT CCT CAT GAA TGC TAT GCC AAA GTG TTC GAT GAA TTT AAA CCT CTT 1140
361 C   A   A   A   D   P   H   E   C   Y   A   K   V   F   D   E   F   K   P   L   380

1141 GTG GAA GAG CCT CAG AAT TTA ATC AAA CAA AAC TGT GAG CTT TTT GAG CAG CTT GGA GAG 1200
381 V   E   E   P   Q   N   L   I   K   Q   N   C   E   L   F   E   Q   L   G   E   400

1201 TAC AAA TTC CAG AAT GCG CTA TTA GTT CGT TAC ACC AAG AAA GTA CCC CAA GTG TCA ACT 1260
401 Y   K   F   Q   N   A   L   L   V   R   Y   T   K   K   V   P   Q   V   S   T   420

1261 CCA ACT CTT GTA GAG GTC TCA AGA AAC CTA GGA AAA GTG GGC AGC AAA TGT TGT AAA CAT 1320
421 P   T   L   V   E   V   S   R   N   L   G   K   V   G   S   K   C   C   K   H   440

1321 CCT GAA GCA AAA AGA ATG CCC TGT GCA GAA GAC TAT CTA TCC GTG GTC CTG AAC CAG TTA 1380
441 P   E   A   K   R   M   P   C   A   E   D   Y   L   S   V   V   L   N   Q   L   460

1381 TGT GTG TTG CAT GAG AAA ACG CCA GTA AGT GAC AGA GTC ACA AAA TGC TGC ACA GAG TCC 1440
461 C   V   L   H   E   K   T   P   V   S   D   R   V   T   K   C   C   T   E   S   480
```

Figure 15C

```
1441 TTG GTG AAC AGG CGA CCA TGC TTT TCA GCT CTG GAA GTC GAT GAA ACA TAC GTT CCC AAA 1500
 481 L   V   N   R   R   P   C   F   S   A   L   E   V   D   E   T   Y   V   P   K   500

1501 GAG TTT AAT GCT GAA ACA TTC ACC TTC CAT GCA GAT ATA TGC ACA CTT TCT GAG AAG GAG 1560
 501 E   F   N   A   E   T   F   T   F   H   A   D   I   C   T   L   S   E   K   E   520

1561 AGA CAA ATC AAG AAA CAA ACT GCA CTT GTT GAG CTT GTG AAA CAC AAG CCC AAG GCA ACA 1620
 521 R   Q   I   K   K   Q   T   A   L   V   E   L   V   K   H   K   P   K   A   T   540

1621 AAA GAG CAA CTG AAA GCT GTT ATG GAT GAT TTC GCA GCT TTT GTA GAG AAG TGC TGC AAG 1680
 541 K   E   Q   L   K   A   V   M   D   D   F   A   A   F   V   E   K   C   C   K   560

1681 GCT GAC GAT AAG GAG ACC TGC TTT GCC GAG GAG GGT AAA AAA CTT GTT GCT GCA AGT CAA 1740
 561 A   D   D   K   E   T   C   F   A   E   E   G   K   K   L   V   A   A   S   Q   580

1741 GCT GCC TTA GGC TTA TAA CAT CTA CAT TTA AAA GCA TCT CAG 1782
 581 A   A   L   G   L   *                                  585
```

Figure 15D

SEQUENCE LISTING

```
<110> Delta Biotechnology Limited
      Principia Pharmaceutical Corporation

<120> Albumin Fusion Proteins

<130> PF542PCT

<140> Unassigned
<141> 2001-04-12

<150> 60/229,358
<151> 2000-04-12

<150> 60/256,931
<151> 2000-12-21

<150> 60/199,384
<151> 2000-04-25

<160> 37

<170> PatentIn Ver. 2.1

<210> 1
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<221> primer_bind
<223> primer useful to clone human growth hormone cDNA

<400> 1
cccaagaatt cccttatcca ggc                                           23


<210> 2
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<221> primer_bind
<223> primer useful to clone human growth hormone cDNA

<400> 2
gggaagctta gaagccacag gatccctcca cag                                33


<210> 3
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
```

<221> misc_structure
<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 3
gataaagatt cccaac 16

<210> 4
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_structure
<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 4
aattgttggg aatcttt 17

<210> 5
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_structure
<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 5
ttaggcttat tcccaac 17

<210> 6
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_structure
<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 6
aattgttggg aataagcc 18

<210> 7
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<221> SITE
<222> 1)..(19)
<223> invertase leader sequence

<220>
<221> SITE
<222> 20)..(24)
<223> first 5 amino acids of mature human serum albumin

<400> 7
Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys
1 5 10 15

Ile Ser Ala Asp Ala His Lys Ser
20

<210> 8
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_structure
<223> synthetic oligonucleotide used to join DNA
fragments with non-cohesive ends.

<400> 8
gagatgcaca cctgagtgag g 21

<210> 9
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_structure
<223> synthetic oligonucleotide used to join DNA
fragments with non-cohesive ends.

<400> 9
gatcctgtgg cttcgatgca cacaaga 27

<210> 10
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_structure
<223> synthetic oligonucleotide used to join DNA
fragments with non-cohesive ends.

<400> 10
ctcttgtgtg catcgaagcc acag 24

<210> 11
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_structure
<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 11
tgtggaagag cctcagaatt tattcccaac 30

<210> 12
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_structure
<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 12
aattgttggg aataaattct gaggctcttc c 31

<210> 13
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_structure
<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 13
ttaggcttag gtggcggtgg atccggcggt ggtggatctt tcccaac 47

<210> 14
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_structure
<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 14
aattgttggg aaagatccac caccgccgga tccaccgcca cctaagcc 48

<210> 15
<211> 62
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_structure
<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 15
ttaggcttag gcggtggtgg atctggtggc ggcggatctg gtggcggtgg atccttccca 60
ac 62

<210> 16
<211> 63
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_structure
<223> synthetic oligonucleotide used to join DNA
fragments with non-cohesive ends.

<400> 16
aattgttggg aaggatccac cgccaccaga tccgccgcca ccagatccac caccgcctaa 60
gcc 63

<210> 17
<211> 1782
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1755)

<400> 17
gat gca cac aag agt gag gtt gct cat cgg ttt aaa gat ttg gga gaa 48
Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
1 5 10 15

gaa aat ttc aaa gcc ttg gtg ttg att gcc ttt gct cag tat ctt cag 96
Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
20 25 30

cag tgt cca ttt gaa gat cat gta aaa tta gtg aat gaa gta act gaa 144
Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
35 40 45

ttt gca aaa aca tgt gtt gct gat gag tca gct gaa aat tgt gac aaa 192
Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
50 55 60

tca ctt cat acc ctt ttt gga gac aaa tta tgc aca gtt gca act ctt 240
Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
65 70 75 80

cgt gaa acc tat ggt gaa atg gct gac tgc tgt gca aaa caa gaa cct 288
Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
85 90 95

gag aga aat gaa tgc ttc ttg caa cac aaa gat gac aac cca aac ctc 336
Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
100 105 110

ccc cga ttg gtg aga cca gag gtt gat gtg atg tgc act gct ttt cat 384
Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
115 120 125

gac aat gaa gag aca ttt ttg aaa aaa tac tta tat gaa att gcc aga 432
Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
130 135 140

aga cat cct tac ttt tat gcc ccg gaa ctc ctt ttc ttt gct aaa agg 480
Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
145 150 155 160

tat aaa gct gct ttt aca gaa tgt tgc caa gct gct gat aaa gct gcc 528
Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
165 170 175

tgc ctg ttg cca aag ctc gat gaa ctt cgg gat gaa ggg aag gct tcg 576
Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
180 185 190

tct gcc aaa cag aga ctc aaa tgt gcc agt ctc caa aaa ttt gga gaa 624
Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
195 200 205

aga gct ttc aaa gca tgg gca gtg gct cgc ctg agc cag aga ttt ccc 672
Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro
210 215 220

aaa gct gag ttt gca gaa gtt tcc aag tta gtg aca gat ctt acc aaa 720
Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys
225 230 235 240

gtc cac acg gaa tgc tgc cat gga gat ctg ctt gaa tgt gct gat gac 768
Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp
245 250 255

agg gcg gac ctt gcc aag tat atc tgt gaa aat cag gat tcg atc tcc 816
Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser
260 265 270

agt aaa ctg aag gaa tgc tgt gaa aaa cct ctg ttg gaa aaa tcc cac 864
Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His
275 280 285

tgc att gcc gaa gtg gaa aat gat gag atg cct gct gac ttg cct tca 912
Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser
290 295 300

tta gct gct gat ttt gtt gaa agt aag gat gtt tgc aaa aac tat gct 960
Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala
305 310 315 320

gag gca aag gat gtc ttc ctg ggc atg ttt ttg tat gaa tat gca aga 1008
Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg
325 330 335

agg cat cct gat tac tct gtc gtg ctg ctg ctg aga ctt gcc aag aca 1056
Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr
340 345 350

```
tat gaa acc act cta gag aag tgc tgt gcc gct gca gat cct cat gaa    1104
Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu
        355                 360                 365

tgc tat gcc aaa gtg ttc gat gaa ttt aaa cct ctt gtg gaa gag cct    1152
Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro
    370                 375                 380

cag aat tta atc aaa caa aac tgt gag ctt ttt gag cag ctt gga gag    1200
Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu
385                 390                 395                 400

tac aaa ttc cag aat gcg cta tta gtt cgt tac acc aag aaa gta ccc    1248
Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro
                405                 410                 415

caa gtg tca act cca act ctt gta gag gtc tca aga aac cta gga aaa    1296
Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys
            420                 425                 430

gtg ggc agc aaa tgt tgt aaa cat cct gaa gca aaa aga atg ccc tgt    1344
Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys
        435                 440                 445

gca gaa gac tat cta tcc gtg gtc ctg aac cag tta tgt gtg ttg cat    1392
Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His
    450                 455                 460

gag aaa acg cca gta agt gac aga gtc aca aaa tgc tgc aca gag tcc    1440
Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
465                 470                 475                 480

ttg gtg aac agg cga cca tgc ttt tca gct ctg gaa gtc gat gaa aca    1488
Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr
                485                 490                 495

tac gtt ccc aaa gag ttt aat gct gaa aca ttc acc ttc cat gca gat    1536
Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
            500                 505                 510

ata tgc aca ctt tct gag aag gag aga caa atc aag aaa caa act gca    1584
Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
        515                 520                 525

ctt gtt gag ctt gtg aaa cac aag ccc aag gca aca aaa gag caa ctg    1632
Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
    530                 535                 540

aaa gct gtt atg gat gat ttc gca gct ttt gta gag aag tgc tgc aag    1680
Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
545                 550                 555                 560

gct gac gat aag gag acc tgc ttt gcc gag gag ggt aaa aaa ctt gtt    1728
Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
                565                 570                 575

gct gca agt caa gct gcc tta ggc tta taacatctac atttaaaagc atctcag 1782
Ala Ala Ser Gln Ala Ala Leu Gly Leu
            580                 585
```

<210> 18
<211> 585
<212> PRT
<213> Homo Sapiens

<400> 18
Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
1 5 10 15

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
20 25 30

Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
35 40 45

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
50 55 60

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
65 70 75 80

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
85 90 95

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
100 105 110

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
115 120 125

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
130 135 140

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
145 150 155 160

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
165 170 175

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
180 185 190

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
195 200 205

Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro
210 215 220

Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys
225 230 235 240

Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp
245 250 255

Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser
260 265 270

Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His
275 280 285

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser
290 295 300

Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala
305 310 315 320

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg
325 330 335

Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr
340 345 350

Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu
355 360 365

Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro
370 375 380

Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu
385 390 395 400

Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro
405 410 415

Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys
420 425 430

Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys
435 440 445

Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His
450 455 460

Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
465 470 475 480

Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr
485 490 495

Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
500 505 510

Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
515 520 525

Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
530 535 540

Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
545 550 555 560

Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
565 570 575

Ala Ala Ser Gln Ala Ala Leu Gly Leu
580 585

<210> 19
<211> 57
<212> DNA
<213> Artificial Sequence

<220>
<221> primer_bind
<223> primer used to generate XhoI and ClaI
site in pPPC0006

<400> 19
gcctcgagaa aagagatgca cacaagagtg aggttgctca tcgatttaaa gatttgg 57

<210> 20
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<221> primer_bind
<223> primer used in generation XhoI and ClaI
site in pPPC0006

<400> 20
aatcgatgag caacctcact cttgtgtgca tctcttttct cgaggctcct ggaataag 58

<210> 21
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<221> primer_bind
<223> primer used in generation XhoI and ClaI
site in pPPC0006

<400> 21
tacaaactta agagtccaat tagc 24

<210> 22
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<221> primer_bind
<223> primer used in generation XhoI and ClaI
site in pPPC0006

<400> 22
cacttctcta gagtggtttc atatgtctt 29

<210> 23
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<221> Misc_Structure
<223> Synthetic oligonucleotide used to alter restriction
sites in pPPC0007

<400> 23
aagctgcctt aggcttataa taaggcgcgc cggccggccg tttaaactaa gcttaattct 60

<210> 24
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<221> Misc_Structure
<223> Synthetic oligonucleotide used to alter restriction
sites in pPPC0007

<400> 24
agaattaagc ttagtttaaa cggccggccg gcgcgcctta ttataagcct aaggcagctt 60

<210> 25
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<221> primer_bind
<223> forward primer useful for generation of albumin
fusion protein in which the albumin moiety is N-terminal
of the Therapeutic Protein

<220>
<221> misc_feature
<222> (18)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (19)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (20)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (21)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (22)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (23)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (24)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (25)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (26)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (27)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (28)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (29)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (30)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (31)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (32)
<223> n equals a,t,g, or c

<400> 25
aagctgcctt aggcttannn nnnnnnnnnn nn 32

<210> 26
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<221> primer_bind

<223> reverse primer useful for generation of albumin fusion protein in which the albumin moiety is N-terminal of the Therapeutic Protein

<220>
<221> misc_feature
<222> (37)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (38)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (39)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (40)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (41)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (42)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (43)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (44)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (45)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (46)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (47)
<223> n equals a,t,g, or c

<220>

<221> misc_feature
<222> (48)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (49)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (50)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (51)
<223> n equals a,t,g, or c

<400> 26
gcgcgcgttt aaacggccgg ccggcgcgcc ttattannnn nnnnnnnnnn n 51

<210> 27
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> forward primer useful for generation of albumin fusion protein in which the albumin moiety is c-terminal of the Therapeutic Protein

<220>
<221> misc_feature
<222> (19)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (20)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (21)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (22)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (23)
<223> n equals a,t,g, or c

<220>
<221> misc_feature

222> (24)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (25)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (26)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (27)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (28)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (29)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (30)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (31)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
222> (32)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (33)
<223> n equals a,t,g, or c

<400> 27
aggagcgtcg acaaaagann nnnnnnnnnn nnn 33

<210> 28
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<221> primer_bind
<223> reverse primer useful for generation of albumin

fusion protein in which the albumin moiety is c-terminal of the Therapeutic Protein

<220>
<221> misc_feature
<222> (38)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (39)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (40)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (41)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (42)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (43)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (44)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (45)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (46)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (47)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (48)
<223> n equals a,t,g, or c

<220>
<221> misc_feature

<222> (49)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (50)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (51)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (52)
<223> n equals a,t,g, or c

<400> 28
ctttaaatcg atgagcaacc tcactcttgt gtgcatcnnn nnnnnnnnnn nn 52

<210> 29
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<221> signal
<223> signal peptide of natural human serum albumin protein

<400> 29
Met Lys Trp Val Ser Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
1 5 10 15

Tyr Ser Arg Ser Leu Asp Lys Arg
20

<210> 30
<211> 114
<212> DNA
<213> Artificial Sequence

<220>
<221> primer_bind
<223> forward primer useful for generation of PC4:HSA
albumin fusion VECTOR

<220>
<221> misc_feature
<222> (5)..(10)
<223> BamHI retsriction site

<220>
<221> misc_feature
<222> (11)..(16)
<223> Hind III retsriction site

<220>
<221> misc_feature

<222> (17)..(27)
<223> Kozak sequence

<220>
<221> misc_feature
<222> (25)..(97)
<223> cds natural signal sequence of human serum albumin

<220>
<221> misc_feature
<222> (75)..(81)
<223> XhoI restriction site

<220>
<221> misc_feature
<222> (98)..(114)
<223> cds first six amino acids of human serum albumin

<400> 30
tcagggatcc aagcttccgc caccatgaag tgggtaacct ttatttccct tctttttctc 60

tttagctcgg cttactcgag ggtgtgtttt cgtcgagatg cacacaagag tgag 114

<210> 31
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<221> primer_bind
<223> reverse primer useful for generation of
PC4:HSA albumin fusion VECTOR

<220>
<221> misc_feature
<222> (6)..(11)
<223> Asp718 restriction site

<220>
<221> misc_feature
<222> (12)..(17)
<223> EcoRI restriction site

<220>
<221> misc_feature
<222> (15)..(17)
<223> reverse complement of stop codon

<220>
<221> misc_feature
<222> (18)..(25)
<223> AscI restriction site

<220>
<221> misc_feature
<222> (18)..(43)
<223> reverse complement of DNA sequence encoding last 9 amino acids

<400> 31

gcagcggtac cgaattcggc gcgccttata agcctaaggc agc 43

<210> 32
<211> 46
<212> DNA
<213> Artificial Sequence

<220>
<221> primer_bind
<223> *forward primer useful for inserting Therapeutic* protein into pC4:HSA vector

<220>
<221> misc_feature
<222> (29)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (30)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (31)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (32)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (33)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (34)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (35)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (36)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (37)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (38)

<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (39)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (40)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (41)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (42)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (43)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (44)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (45)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (46)
<223> n equals a,t,g, or c

<400> 32
ccgccgctcg aggggtgtgt ttcgtcgann nnnnnnnnnn nnnnnn 46

<210> 33
<211> 55
<212> DNA
<213> Artificial Sequence

<220>
<221> primer_bind
<223> reverse primer useful for inserting Therapeutic
protein into pC4:HSA vector

<220>
<221> misc_feature
<222> (38)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (39)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (40)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (41)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (42)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (43)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (44)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (45)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (46)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (47)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (48)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (49)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (50)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (51)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (52)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (53)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (54)
<223> n equals a,t,g, or c

<220>
<221> misc_feature
<222> (55)
<223> n equals a,t,g, or c

<400> 33
agtcccatcg atgagcaacc tcactcttgt gtgcatcnnn nnnnnnnnnn nnnnn 55

<210> 34
<211> 17
<212> PRT
<213> Artificial Sequence

<220>
<221> signal
<223> Stanniocalcin signal peptide

<400> 34
Met Leu Gln Asn Ser Ala Val Leu Leu Leu Leu Val Ile Ser Ala Ser
1 5 10 15

Ala

<210> 35
<211> 22
<212> PRT
<213> Artificial Sequence

<220>
<221> signal
<223> Synthetic signal peptide

<400> 35
Met Pro Thr Trp Ala Trp Trp Leu Phe Leu Val Leu Leu Leu Ala Leu
1 5 10 15

Trp Ala Pro Ala Arg Gly
20

<210> 36

<211> 66
<212> PRT
<213> Agkistrodon piscivorus

<400> 36
Ile Thr Tyr Thr Asp Cys Thr Glu Ser Gly Gln Asn Leu Cys Leu Cys
1 5 10 15

Glu Gly Ser Asn Val Cys Gly Lys Gly Asn Lys Cys Ile Leu Gly Ser
20 25 30

Gln Gly Lys Asp Asn Gln Cys Val Thr Gly Glu Gly Thr Pro Lys Pro
35 40 45

Gln Ser His Asn Gln Gly Asp Phe Glu Pro Ile Pro Glu Asp Ala Tyr
50 55 60

Asp Glu
65

<210> 37
<211> 71
<212> PRT
<213> Agkistrodon piscivorus

<400> 37
Glu Ala Gly Glu Glu Cys Asp Cys Gly Ser Pro Glu Asn Pro Cys Cys
1 5 10 15

Asp Ala Ala Thr Cys Lys Leu Arg Pro Gly Ala Gln Cys Ala Glu Gly
20 25 30

Leu Cys Cys Asp Gln Cys Lys Phe Met Lys Glu Gly Thr Val Cys Arg
35 40 45

Ala Arg Gly Asp Asp Val Asn Asp Tyr Cys Asn Gly Ile Ser Ala Gly
50 55 60

Cys Pro Arg Asn Pro Phe His
65 70

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/12009

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) :C07K 14/00; C07H 21/04; C12N 15/00; A61K 38/16
US CL : 435/320.1; 514/2; 530/350; 536/23.4
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/320.1; 514/2; 530/350; 536/23.4

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
MEDLINE, BIOSIS, EMBASE, CAPLUS, various sequence databases
search terms: chimera, fusion, human serum albumin, interferon, cytokine, erythropoietin, SEQ ID No. 18

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X — Y	WO 99/66054 A2 (GENZYME TRANSGENICS CORP.) 23 December 1999, see entire document.	1,2,4,6,9,10,12,14,16,17,24,28,30-37,40-42 — 3,15,25;26,29,39N
X — Y	WO 97/24445 A1 (DELTA BIOTECHNOLOGY LIMITED) 10 July 1997, seen entire document, especially pp. 3-6, 12, 13, and 19-26	1-4,6,9,10,13-17,25,28-34,36,37,39-42 — 5,7,8,24,26,27,35,38

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date		
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search 16 JULY 2001	Date of mailing of the international search report 22 AUG 2001
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer TERRY J. DEY BRONWEN M. LOEB PARALEGAL SPECIALIST Telephone No. (703) TECHNOLOGY CENTER 1600

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/12009

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,705,363 A (IMAKAWA, K.) 6 January 1998, see entire document, especially col. 23, line 44-col. 24, line 56.	1,2,4,5-7,10,25,33-38,40-42
Y		3,14,15,16,17,24,
X	US 5,766,883 A (BALLANCE et al) 16 June 1998, see entire document, especially col.1 ,line 38-54, col. 6, line 53-col 10, line 41, col. 11, line 36-col. 12, line 40.	1,2,4,6,10,13,14,25,30-34,36,37,40-42

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 October 2001 (25.10.2001)

PCT

(10) International Publication Number
WO 01/79271 A1

(51) International Patent Classification[7]: C07K 14/00, C07H 21/04, C12N 15/00, A61K 38/16

(21) International Application Number: PCT/US01/12009

(22) International Filing Date: 12 April 2001 (12.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/229,358	12 April 2000 (12.04.2000)	US
60/199,384	25 April 2000 (25.04.2000)	US
60/256,931	21 December 2000 (21.12.2000)	US

(71) Applicants *(for all designated States except US)*: **PRINCIPIA PHARMACEUTICAL CORPORATION** [US/US]; 2650 Eisenhower Avenue, Building C, Norristown, PA 19403 (US). **DELTA BIOTECHNOLOGY LIMITED** [GB/GB]; Castle Court, 59 Castle Boulevard, Nottingham NG7 1FD (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **BALLANCE, David, James** [GB/US]; 1113 Cymry Drive, Berwyn, PA 19312 (US). **SLEEP, Darrell** [GB/US]; 66 Ladybay Road, West Bridgford, Nottingham NG2 5DS (GB). **TURNER, Andrew, John** [GB/US]; Apartment C-28, 305 Conestoga Way, Eagleville, PA 19408 (US). **SADEGHI, Homayoun** [US/US]; 320 E. Court Street, Doylestown, PA 18901 (US). **PRIOR, Christopher, P.** [US/US]; 460 Wyldhaven Road, Rosemont, PA 19010 (US).

(74) Agents: **GARRETT, Arthur, S.** et al.; Finnegan, Henderson, Farabow, Garrett & Dunner, LLP, 1300 I Street, NW., Washington, DC 20005-3315 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *with (an) indication(s) in relation to deposited biological material furnished under Rule 13bis separately from the description*

(48) Date of publication of this corrected version: 28 February 2002

(15) Information about Correction: see PCT Gazette No. 09/2002 of 28 February 2002, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/79271 A1

(54) Title: ALBUMIN FUSION PROTEINS

(57) Abstract: The present invention encompasses albumin fusion proteins. Nucleic acid molecules encoding the albumin fusion proteins of the invention are also encompassed by the invention, as are vectors containing these nucleic acids, host cells transformed with these nucleic acids vectors, and methods of making the albumin fusion proteins of the invention and using these nucleic acids, vectors, and/or host cells. Additionally the present invention encompasses pharmaceutical compositions comprising albumin fusion proteins and methods of treating, preventing, or ameliorating diseases, disorders or conditions using albumin fusion proteins of the invention.

ALBUMIN FUSION PROTEINS

BACKGROUND OF THE INVENTION

The invention relates generally to Therapeutic proteins (including, but not limited to, a polypeptide, antibody, or peptide, or fragments and variants thereof) fused to albumin or fragments or variants of albumin. The invention further relates to Therapeutic proteins (including, but not limited to, a polypeptide, antibody, or peptide, or fragments and variants thereof) fused to albumin or fragments or variants of albumin, that exhibit extended shelf-life and/or extended or therapeutic activity in solution. These fusion proteins are herein collectively referred to as "albumin fusion proteins of the invention." The invention encompasses therapeutic albumin fusion proteins, compositions, pharmaceutical compositions, formulations and kits. Nucleic acid molecules encoding the albumin fusion proteins of the invention are also encompassed by the invention, as are vectors containing these nucleic acids, host cells transformed with these nucleic acids vectors, and methods of making the albumin fusion proteins of the invention using these nucleic acids, vectors, and/or host cells.

The invention is also directed to methods of *in vitro* stabilizing a Therapeutic protein via fusion or conjugation of the Therapeutic protein to albumin or fragments or variants of albumin.

Human serum albumin (HSA, or HA), a protein of 585 amino acids in its mature form (as shown in Figure 15 or in SEQ ID NO:18), is responsible for a significant proportion of the osmotic pressure of serum and also functions as a carrier of endogenous and exogenous ligands. At present, HA for clinical use is produced by extraction from human blood. The production of recombinant HA (rHA) in microorganisms has been disclosed in EP 330 451 and EP 361 991.

The role of albumin as a carrier molecule and its inert nature are desirable properties for use as a carrier and transporter of polypeptides *in vivo*. The use of albumin as a component of an albumin fusion protein as a carrier for various proteins has been suggested in WO 93/15199, WO 93/15200, and EP 413 622. The use of N-terminal fragments of HA

Therapeutic protein may be achieved by genetic manipulation, such that the DNA coding for HA, or a fragment thereof, is joined to the DNA coding for the Therapeutic protein. A suitable host is then transformed or transfected with the fused nucleotide sequences, so arranged on a suitable plasmid as to express a fusion polypeptide. The expression may be effected *in vitro* from, for example, prokaryotic or eukaryotic cells, or *in vivo e.g.* from a transgenic organism.

Therapeutic proteins in their native state or when recombinantly produced, such as interferons and growth hormones, are typically labile molecules exhibiting short shelf-lives, particularly when formulated in aqueous solutions. The instability in these molecules when formulated for administration dictates that many of the molecules must be lyophilized and refrigerated at all times during storage, thereby rendering the molecules difficult to transport and/or store. Storage problems are particularly acute when pharmaceutical formulations must be stored and dispensed outside of the hospital environment. Many protein and peptide drugs also require the addition of high concentrations of other protein such as albumin to reduce or prevent loss of protein due to binding to the container. This is a major concern with respect to proteins such as IFN. For this reason, many Therapeutic proteins are formulated in combination with large proportion of albumin carrier molecule (100-1000 fold excess), though this is an undesirable and expensive feature of the formulation.

Few practical solutions to the storage problems of labile protein molecules have been proposed. Accordingly, there is a need for stabilized, long lasting formulations of proteinaceous therapeutic molecules that are easily dispensed, preferably with a simple formulation requiring minimal post-storage manipulation.

SUMMARY OF THE INVENTION

The present invention is based, in part, on the discovery that Therapeutic proteins may be stabilized to extend the shelf-life, and/or to retain the Therapeutic protein's activity for extended periods of time in solution, *in vitro* and/or *in vivo*, by genetically or chemically fusing or conjugating the Therapeutic protein to albumin or a fragment (portion) or variant of albumin, that is sufficient to stabilize the protein and/or its activity. In addition it has been determined that the use of albumin-fusion proteins or albumin conjugated proteins may reduce the need to formulate protein solutions with large excesses of carrier proteins (such as albumin, unfused) to prevent loss of Therapeutic proteins due to factors such as binding to the container.

The present invention encompasses albumin fusion proteins comprising a Therapeutic protein (e.g., a polypeptide, antibody, or peptide, or fragments and variants thereof) fused to albumin or a fragment (portion) or variant of albumin.. The present invention also encompasses albumin fusion proteins comprising a Therapeutic protein (e.g., a polypeptide,

antibody, or peptide, or fragments and variants thereof) fused to albumin or a fragment (portion) or variant of albumin, that is sufficient to prolong the shelf life of the Therapeutic protein, and/or stabilize the Therapeutic protein and/or its activity in solution (or in a pharmaceutical composition) *in vitro* and/or *in vivo*. Nucleic acid molecules encoding the albumin fusion proteins of the invention are also encompassed by the invention, as are vectors containing these nucleic acids, host cells transformed with these nucleic acids vectors, and methods of making the albumin fusion proteins of the invention and using these nucleic acids, vectors, and/or host cells.

The invention also encompasses pharmaceutical formulations comprising an albumin fusion protein of the invention and a pharmaceutically acceptable diluent or carrier. Such formulations may be in a kit or container. Such kit or container may be packaged with instructions pertaining to the extended shelf life of the Therapeutic protein. Such formulations may be used in methods of treating, preventing, ameliotationg or diagnosing a disease or disease symptom in a patient, preferably a mammal, most preferably a human, comprising the step of administering the pharmaceutical formulation to the patient.

In other embodiments, the present invention encompasses methods of preventing treating, or ameliorating a disease or disorder. In preferred embodiments, the present invention encompasses a method of treating a disease or disorder listed in the "Preferred Indication Y" column of Table 1 comprising administering to a patient in which such treatment, prevention or amelioration is desired an albumin fusion protein of the invention that comprises a Therapeutic protein portion corresponding to a Therapeutic protein (or fragment or variant thereof) disclosed in the "Therapeutic Protein X" column of Table 1 (in the same row as the disease or disorder to be treated is listed in the "Preferred Indication Y" column of Table 1) in an amount effective to treat prevent or ameliorate the disease or disorder.

In another embodiment, the invention includes a method of extending the shelf life of a Therapeutic protein (e.g., a polypeptide, antibody, or peptide, or fragments and variants thereof) comprising the step of fusing or conjugating the Therapeutic protein to albumin or a fragment (portion) or variant of albumin, that is sufficient to extend the shelf-life of the Therapeutic protein. In a preferred embodiment, the Therapeutic protein used according to this method is fused to the albumin, or the fragment or variant of albumin. In a most preferred embodiment, the Therapeutic protein used according to this method is fused to albumin, or a fragment or variant of albumin, via recombinant DNA technology or genetic engineering.

In another embodiment, the invention includes a method of stabilizing a Therapeutic protein (e.g., a polypeptide, antibody, or peptide, or fragments and variants thereof) in solution, comprising the step of fusing or conjugating the Therapeutic protein to albumin or a fragment (portion) or variant of albumin, that is sufficient to stabilize the Therapeutic protein.

In a preferred embodiment, the Therapeutic protein used according to this method is fused to the albumin, or the fragment or variant of albumin. In a most preferred embodiment, the Therapeutic protein used according to this method is fused to albumin, or a fragment or variant of albumin, via recombinant DNA technology or genetic engineering.

The present invention further includes transgenic organisms modified to contain the nucleic acid molecules of the invention, preferably modified to express the albumin fusion proteins encoded by the nucleic acid molecules.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 depicts the extended shelf-life of an HA fusion protein in terms of the biological activity (Nb2 cell proliferation) of HA-hGH remaining after incubation in cell culture media for up to 5 weeks at 37°C. Under these conditions, hGH has no observed activity by week 2.

Figure 2 depicts the extended shelf-life of an HA fusion protein in terms of the stable biological activity (Nb2 cell proliferation) of HA-hGH remaining after incubation in cell culture media for up to 3 weeks at 4, 37, or 50°C. Data is normalized to the biological activity of hGH at time zero.

Figures 3A and 3B compare the biological activity of HA-hGH with hGH in the Nb2 cell proliferation assay. Figure 3A shows proliferation after 24 hours of incubation with various concentrations of hGH or the albumin fusion protein, and Figure 3B shows proliferation after 48 hours of incubation with various concentrations of hGH or the albumin fusion protein.

Figure 4 shows a map of a plasmid (pPPC0005) that can be used as the base vector into which polynucleotides encoding the Therapeutic proteins (including polypeptides and fragments and variants thereof) may be cloned to form HA-fusions. Plasmid Map key: PRB1p: *PRB1 S. cerevisiae* promoter; FL: Fusion leader sequence; rHA: cDNA encoding HA; ADH1t: *ADH1 S. cerevisiae* terminator; T3: T3 sequencing primer site; T7: T7 sequencing primer site; Amp R: β-lactamase gene; ori: origin of replication. Please note that in the provisional applications to which this application claims priority, the plasmid in Figure 4 was labeled pPPC0006, instead of pPPC0005. In addition the drawing of this plasmid did not show certain pertinent restriction sites in this vector. Thus in the present application, the drawing is labeled pPPC0005 and more restriction sites of the same vector are shown.

Figure 5 compares the recovery of vial-stored HA-IFN solutions of various concentrations with a stock solution after 48 or 72 hours of storage.

Figure 6 compares the activity of an HA-α-IFN fusion protein after administration to monkeys via IV or SC.

Figure 7 describes the bioavailability and stability of an HA-α-IFN fusion protein.

Figure 8 is a map of an expression vector for the production of HA-α-IFN.

Figure 9 shows the location of loops in HA.

Figure 10 is an example of the modification of an HA loop.

Figure 11 is a representation of the HA loops.

Figure 12 shows the HA loop IV.

Figure 13 shows the tertiary structure of HA.

Figure 14 shows an example of a scFv-HA fusion

Figure 15 shows the amino acid sequence of the mature form of human albumin (SEQ ID NO:18) and a polynucleotide encoding it (SEQ ID NO:17).

DETAILED DESCRIPTION

As described above, the present invention is based, in part, on the discovery that a Therapeutic protein (e.g., a polypeptide, antibody, or peptide, or fragments and variants thereof) may be stabilized to extend the shelf-life and/or retain the Therapeutic protein's activity for extended periods of time in solution (or in a pharmaceutical composition) *in vitro* and/or *in vivo*, by genetically fusing or chemically conjugating the Therapeutic protein, polypeptide or peptide to all or a portion of albumin sufficient to stabilize the protein and its activity.

The present invention relates generally to albumin fusion proteins and methods of treating, preventing, or ameliorating diseases or disorders. As used herein, "albumin fusion protein" refers to a protein formed by the fusion of at least one molecule of albumin (or a fragment or variant thereof) to at least one molecule of a Therapeutic protein (or fragment or variant thereof). An albumin fusion protein of the invention comprises at least a fragment or variant of a Therapeutic protein and at least a fragment or variant of human serum albumin, which are associated with one another, preferably by genetic fusion (i.e., the albumin fusion protein is generated by translation of a nucleic acid in which a polynucleotide encoding all or a portion of a Therapeutic protein is joined in-frame with a polynucleotide encoding all or a portion of albumin) or chemical conjugation to one another. The Therapeutic protein and albumin protein, once part of the albumin fusion protein, may be referred to as a "portion", "region" or "moiety" of the albumin fusion protein (e.g., a "Therapeutic protein portion" or an "albumin protein portion").

In one embodiment, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a Therapeutic protein (e.g., as described in Table 1) and a serum albumin protein. In other embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a biologically active and/or therapeutically active fragment of a Therapeutic protein and a serum albumin protein. In other embodiments, the

invention provides an albumin fusion protein comprising, or alternatively consisting of, a biologically active and/or therapeutically active variant of a Therapeutic protein and a serum albumin protein. In preferred embodiments, the serum albumin protein component of the albumin fusion protein is the mature portion of serum albumin.

In further embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a Therapeutic protein, and a biologically active and/or therapeutically active fragment of serum albumin. In further embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a Therapeutic protein and a biologically active and/or therapeutically active variant of serum albumin. In preferred embodiments, the Therapeutic protein portion of the albumin fusion protein is the mature portion of the Therapeutic protein. In a further preferred embodiment, the Therapeutic protein portion of the albumin fusion protein is the extracellular soluble domain of the Therapeutic protein. In an alternative embodiment, the Therapeutic protein portion of the albumin fusion protein is the active form of the Therapeutic protien.

In further embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a biologically active and/or therapeutically active fragment or variant of a Therapeutic protein and a biologically active and/or therapeutically active fragment or variant of serum albumin. In preferred embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, the mature portion of a Therapeutic protein and the mature portion of serum albumin.

Therapeutic proteins

As stated above, an albumin fusion protein of the invention comprises at least a fragment or variant of a Therapeutic protein and at least a fragment or variant of human serum albumin, which are associated with one another, preferably by genetic fusion or chemical conjugation.

As used herein, "Therapeutic protein" refers to proteins, polypeptides, antibodies, peptides or fragments or variants thereof, having one or more therapeutic and/or biological activities. Therapeutic proteins encompassed by the invention include but are not limited to, proteins, polypeptides, peptides, antibodies, and biologics. (The terms peptides, proteins, and polypeptides are used interchangeably herein.) It is specifically contemplated that the term "Therapeutic protein" encompasses antibodies and fragments and variants thereof. Thus an albumin fusion protein of the invention may contain at least a fragment or variant of a Therapeutic protein, and/or at least a fragment or variant of an antibody. Additionally, the term "Therapeutic protein" may refer to the endogenous or naturally occurring correlate of a Therapeutic protein.

By a polypeptide displaying a "therapeutic activity" or a protein that is "therapeutically

active" is meant a polypeptide that possesses one or more known biological and/or therapeutic activities associated with a Therapeutic protein such as one or more of the Therapeutic proteins described herein or otherwise known in the art. As a non-limiting example, a "Therapeutic protein" is a protein that is useful to treat, prevent or ameliorate a disease, condition or disorder. As a non-limiting example, a "Therapeutic protein" may be one that binds specifically to a particular cell type (normal (e.g., lymphocytes) or abnormal e.g., (cancer cells)) and therefore may be used to target a compound (drug, or cytotoxic agent) to that cell type specifically.

In another non-limiting example, a "Therapeutic protein" is a protein that has a biological activity, and in particular, a biological activity that is useful for treating preventing or ameliorating a disease. A non-inclusive list of biological activities that may be possessed by a Therapeutic protein includes, enhancing the immune response, promoting angiogenesis, inhibiting angiogenesis, regulating hematopoietic functions, stimulating nerve growth, enhancing an immune response, inhibiting an immune response, or any one or more of the biological activities described in the "Biological Activities" section below.

As used herein, "therapeutic activity" or "activity" may refer to an activity whose effect is consistent with a desirable therapeutic outcome in humans, or to desired effects in non-human mammals or in other species or organisms. Therapeutic activity may be measured *in vivo* or *in vitro*. For example, a desirable effect may be assayed in cell culture. As an example, when hGH is the Therapeutic protein, the effects of hGH on cell proliferation as described in Example 1 may be used as the endpoint for which therapeutic activity is measured. Such *in vitro* or cell culture assays are commonly available for many Therapeutic proteins as described in the art.

Examples of useful assays for particular Therapeutic proteins include, but are not limited to, GMCSF (Eaves, A.C. and Eaves C.J., Erythropoiesis in culture. In: McCullock EA (edt) Cell culture techniques - Clinics in hematology. WB Saunders, Eastbourne, pp 371-91 (1984); Metcalf, D., International Journal of Cell Cloning 10: 116-25 (1992); Testa, N.G., et al., Assays for hematopoietic growth factors. In: Balkwill FR (edt) Cytokines, A practical Approach, pp 229-44; IRL Press Oxford 1991) EPO (bioassay: Kitamura *et al.*, J. Cell. Physiol. 140 p323 (1989)); Hirudin (platelet aggregation assay: Blood Coagul Fibrinolysis 7(2):259-61 (1996)); IFNα (anti-viral assay: Rubinstein et al., J. Virol. 37(2):755-8 (1981); anti-proliferative assay: Gao Y, et al Mol Cell Biol. 19(11):7305-13 (1999); and bioassay: Czarniecki *et al.*, J. Virol. 49 p490 (1984)); GCSF (bioassay: Shirafuji *et al.*, Exp. Hematol. 17 p116 (1989); proliferation of murine NFS-60 cells (Weinstein et al, Proc Natl Acad Sci 83:5010-4 (1986)); insulin (^{3}H-glucose uptake assay: Steppan et al., Nature 409(6818):307-12 (2001)); hGH (Ba/F3-hGHR proliferation assay: J Clin Endocrinol Metab 85(11):4274-9 (2000); International standard for growth hormone: Horm Res, 51

Suppl 1:7-12 (1999)); factor X (factor X activity assay: Van Wijk et al. Thromb Res 22:681-686 (1981)); factor VII (coagulation assay using prothrombin clotting time: Belaaouaj et al., J. Biol. Chem. 275:27123-8(2000); Diaz-Collier et al., Thromb Haemost 71:339-46 (1994)), or as shown in Table 1 in the "Exemplary Activity Assay" column.

Therapeutic proteins corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention, such as cell surface and secretory proteins, are often modified by the attachment of one or more oligosaccharide groups. The modification, referred to as glycosylation, can dramatically affect the physical properties of proteins and can be important in protein stability, secretion, and localization. Glycosylation occurs at specific locations along the polypeptide backbone. There are usually two major types of glycosylation: glycosylation characterized by O-linked oligosaccharides, which are attached to serine or threonine residues; and glycosylation characterized by N-linked oligosaccharides, which are attached to asparagine residues in an Asn-X-Ser/Thr sequence, where X can be any amino acid except proline. N-acetylneuramic acid (also known as sialic acid) is usually the terminal residue of both N-linked and 0-linked oligosaccharides. Variables such as protein structure and cell type influence the number and nature of the carbohydrate units within the chains at different glycosylation sites. Glycosylation isomers are also common at the same site within a given cell type.

For example, several types of human interferon are glycosylated. Natural human interferon-α2 is O-glycosylated at threonine 106, and N-glycosylation occurs at asparagine 72 in interferon-α14 (Adolf *et al.*, J. Biochem 276:511 (1991); Nyman TA *et al.*, J. Biochem 329:295 (1998)). The oligosaccharides at asparagine 80 in natural interferon-$\beta 1\alpha$ may play an important factor in the solubility and stability of the protein, but may not be essential for its biological activity. This permits the production of an unglycosylated analog (interferon-β1b) engineered with sequence modifications to enhance stability (Hosoi *et al.*, J. Interferon Res. 8:375 (1988; Karpusas *et al.*, Cell Mol Life Sci 54:1203 (1998); Knight, J. Interferon Res. 2:421 (1982); Runkel *et al.*, Pharm Res 15:641 (1998); Lin, Dev. Biol. Stand. 96:97 (1998))1. Interferon-γ contains two N-linked oligosaccharide chains at positions 25 and 97, both important for the efficient formation of the bioactive recombinant protein, and having an influence on the pharmacokinetic properties of the protein (Sareneva *et al.*, Eur. J. Biochem 242:191 (1996); Sareneva *et al*,. Biochem J. 303:831 (1994); Sareneva *et al.*, J. Interferon Res. 13:267 (1993)). Mixed O-linked and N-linked glycosylation also occurs, for example in human erythropoietin, N-linked glycosylation occurs at asparagine residues located at positions 24, 38 and 83 while O-linked glycosylation occurs at a serine residue located at

position 126 (Lai *et al.*, J. Biol. Chem. 261:3116 (1986); Broudy *et al.*, Arch. Biochem. Biophys. 265:329 (1988)).

Therapeutic proteins corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention, as well as analogs and variants thereof, may be modified so that glycosylation at one or more sites is altered as a result of manipulation(s) of their nucleic acid sequence, by the host cell in which they are expressed, or due to other conditions of their expression. For example, glycosylation isomers may be produced by abolishing or introducing glycosylation sites, *e.g.*, by substitution or deletion of amino acid residues, such as substitution of glutamine for asparagine, or unglycosylated recombinant proteins may be produced by expressing the proteins in host cells that will not glycosylate them, *e.g.* in *E. coli* or glycosylation-deficient yeast. These approaches are described in more detail below and are known in the art.

Therapeutic proteins corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention include, but are not limited to, plasma proteins. More specifically, such Therapeutic proteins include, but are not limited to, immunoglobulins, serum cholinesterase, alpha-1 antitrypsin, aprotinin, coagulation factors in both pre and active forms including but not limited to, von Willebrand factor, fibrinogen, factor II, factor VII, factor VIIA activated factor, factor VIII, factor IX, factor X, factor XIII, c1 inactivator, antithrombin III, thrombin, prothrombin, apo-lipoprotein, c-reactive protein, and protein C. Therapeutic proteins corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention further include, but are not limited to, human growth hormone (hGH), α-interferon, erythropoietin (EPO), granulocyte-colony stimulating factor (GCSF), granulocyte-macrophage colony-stimulating factor (GMCSF), insulin, single chain antibodies, autocrine motility factor, scatter factor, laminin, hirudin, applaggin, monocyte chemotactic protein (MCP/MCAF), macrophage colony-stimulating factor (M-CSF), osteopontin, platelet factor 4, tenascin, vitronectin, in addition to those described in Table 1. These proteins and nucleic acid sequences encoding these proteins are well known and available in public databases such as Chemical Abstracts Services Databases (e.g., the CAS Registry), GenBank, and GenSeq as shown in Table 1.

Additional Therapeutic proteins corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention include, but are not limited to, one or more of the Therapeutic proteins or peptides disclosed in the "Therapeutic Protein X" column of Table 1, or fragment or variable thereof.

Table 1 provides a non-exhaustive list of Therapeutic proteins that correspond to a Therapeutic protein portion of an albumin fusion protein of the invention. The "Therapeutic Protein X" column discloses Therapeutic protein molecules followed by parentheses containing scientific and brand names that comprise, or alternatively consist of, that

Therapeutic protein molecule or a fragment or variant thereof. "Therapeutic protein X" as used herein may refer either to an individual Therapeutic protein molecule (as defined by the amino acid sequence obtainable from the CAS and Genbank accession numbers), or to the entire group of Therapeutic proteins associated with a given Therapeutic protein molecule disclosed in this column. The "Exemplary Identifier" column provides Chemical Abstracts Services (CAS) Registry Numbers (published by the American Chemical Society) and/or Genbank Accession Numbers ((e.g., Locus ID, NP_XXXXX (Reference Sequence Protein), and XP_XXXXX (Model Protein) identifiers available through the national Center for Biotechnology Information (NCBI) webpage at www.ncbi.nlm.nih.gov) that correspond to entries in the CAS Registry or Genbank database which contain an amino acid sequence of the Therapeutic Protein Molecule or of a fragment or variant of the Therapeutic Protein Molecule. The summary pages associated with each of these CAS and Genbank Accession Numbers are each incorporated by reference in their entireties, particularly with respect to the amino acid sequences described therein. The "PCT/Patent Reference" column provides U.S. Patent numbers, or PCT International Publication Numbers corresponding to patents and/or published patent applications that describe the Therapeutic protein molecule. Each of the patents and/or published patent applications cited in the "PCT/Patent Reference" column are herein incorporated by reference in their entireties. In particular, the amino acid sequences of the specified polypeptide set forth in the sequence listing of each cited "PCT/Patent Reference", the variants of these amino acid sequences (mutations, fragments, etc.) set forth, for example, in the detailed description of each cited "PCT/Patent Reference", the therapeutic indications set forth, for example, in the detailed description of each cited "PCT/Patent Reference", and the activity asssaysfor the specified polypeptide set forth in the detailed description, and more particularly, the examples of each cited "PCT/Patent Reference" are incorporated herein by reference. The "Biological activity" column describes Biological activities associated with the Therapeutic protein molecule. The "Exemplary Activity Assay" column provides references that describe assays which may be used to test the therapeutic and/or biological activity of a Therapeutic protein or an albumin fusion protein of the invention comprising a Therapeutic protein X portion. Each of the references cited in the "Exemplary Activity Assay" column are herein incorporated by reference in their entireties, particularly with respect to the description of the respective activity assay described in the reference (see Methods section, for example) for assaying the corresponding biological activity set forth in the "Biological Activity" column of Table 1. The "Preferred Indication Y" column describes disease, disorders, and/or conditions that may be treated, prevented, diagnosed, or ameliorated by Therapeutic protein X or an albumin fusion protein of the invention comprising a Therapeutic protein X portion.

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Alpha-1-antitrypsin (Alpha-1 proteinase; Alpha-1-trypsin inhibitor; Prolastin; API; API Inhale)	CAS-9041-92-3 LocusID:5265 LocusID:5299 NP_000286 NP_006211 XP_007481 XP_012372	WO8600337 EP103409-A EP155188 US5399684 US5736379 US6025161 US4839283 US4876197-A	Alpha-1-antitrypsin is an an enzyme inhibitor that belongs to the family of serpin serine protease inhibitors. The molecule inhibits the activity of trypsin and elastase.	Enzyme Inhibition asay: Gaillard MC, Kilroe-Smith TA. 1987 Determination of functional activity of alpha 1-protease inhibitor and alpha 2-macroglobulin in human plasma using elastase. J Clin Chem Clin Biochem. 25(3):167-72. Burnouf T, Constans J, Clerc A, Descamps J, Martinache L, Goudemand M. 1987 Biochemical and biological properties of an alpha 1-antitrypsin concentrate. Vox Sang;52(4):291-7	Emphysema; Infant Respiratory Distress Syndrome; Pulmonary Fibrosis; Respiratory Syncytial Virus Infections;Asthma; Cystic Fibrosis; Genitourinary Disorders; HIV Infections Treatment; Inflammatory Bowel Disorders; Skin Disorders; Viral Hepatitis; Alpha-1 Antitrypsin Deficiency; Adult Respiratory Distress Syndrome
Antihemophilic factor (Von Willebrand factor; Von Willebrand factor complex; HELIXATE-FS; HEMOFIL M; KOATE-DVI/HP; ALPHANATE; MONARC-M; HUMATE-P)	LocusID:7450 NP_000543 XP_006947	WO8606096 EP197592 WO9316709-A US5849536 US6008193 US5849702 US5238919	The glycoprotein encoded by this gene H1functions as both an antihemophilic factor carrier and a platelet-vessel wall mediator in the blood coagulation system. It is crucial to the hemostasis process. Mutations in this gene or deficiencies in this protein result in von Willebrand's disease.	Macroscopic platelet agglutination assay (Wright R. Ann Clin Lab Sci 1990 20(1):73). Collagen binding assay for von Willebrand factor (Favaloro EJ. Thromb Haemost 2000 83(1):127)	Hemophilia; Hemophilia A; von Willebrand disease
Antithrombin III (ATIII; Antithrombin-heparin conjugate; ATH; aaATIII (Antithrombin III-modified); Atnativ; Anthrobin; Atenativ; Athimbin; Kybernin; Thrombhibin)	CAS-155319-91-8 CAS-52014-67-2 LocusID:462 NP_000479 XP_001452	WO9100291 EP568833 GB2116183	Serpin serine protease that inhibits thrombin and other proteins involved in blood coagulation	Thrombin activity assay (Verheul et al., Blood 96:4216-4221, 2000)	Sepsis; Thrombosis; Unstable Angina Pectoris; Coagulation disorders; Respiratory Distress Syndrome; Control of blood clotting during coronary artery bypass surgery; Cancer (aaATIII)

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Apo-lipoprotein (Apo E; Apo A4; Apo A1; Apo B)	CAS-150287-52-8 LocusID:335 LocusID:338 LocusID:348 NP_000030 NP_000375 NP_000032 XP_006435 XP_002288 XP_008844	WO8803166-A WO9307165-A US5408038-A. WO9107505-A WO9315198-A US5472858-A US5364793-A WO8702062-A WO9307165-A WO9856938-A1 US4943527-A	ApoA1 promotes cholesterol efflux from tissues to the liver for excretion. ApoA1 is the major protein component of high density lipoprotein (HDL) in the plasma. ApoA1 is a cofactor for lecithin cholesterolacyltransferase (LCAT), which is responsible for the formation of most plasma cholesteryl esters. Defects in the ApoA1 gene are associated with HDL deficiency and Tangier disease. ApoB is the main apolipoprotein of chylomicrons and low density lipoproteins (LDL). ApoB binds triglycerides and clears LDL from circulation. ApoE is a component of lipoproteins and a ligand for low density lipoprotein receptor. ApoE binds to a specific receptor on liver cells and peripheral cells. ApoE is essential for the normal catabolism of triglyceride-rich lipoprotein constituents. Defects in ApoE result in familial dysbetalipoproteinemia, or type III hyperlipoproteinemia (HLP III), in which increased plasma cholesterol and triglycerides are the consequence of impaired clearance of chylomicron and VLDL remnants.	Cholesterol efflux from human fibroblasts can be directly measured in response to lipid reconstituted ApoA1 (J Biol Chem 1996 Oct 11;271(41):25145-51). The capacity of ApoB to participate in LDL clearance can be evaluated by measurements of ApoB binding to hepatic lipase (J Biol Chem, Vol. 273, Issue 32, 20456-20462) and lipoprotein lipase (J Biol Chem. 1995 Apr 7;270(14):8081-6.). ApoE binding to its receptor can be measured directly for example as illustrated for the liver ApoE receptor (J Biol Chem 1986 Mar 25;261(9):4256-67).	Atherosclerosis; Coronary restenosis; Hypercholesterolemia; Hyperlipidemia; Kaposi's Sarcoma

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Applaggin (Agkistrodon piscivorus piscivorus (North American water moccasin snake); disulfide-linked Arg-Gly-Asp-containing dimeric polypeptide)	CAS-129037-76-9 Genbank:A33990	WO9409036-A WO9008772-A WO9210575-A JP05255395-A	Applaggin (Agkistrodon piscivorus piscivorus platelet aggregation inhibitor) is a 17,700-Da polypeptide dimer which a potent inhibitor of platelet activation. Applaggin blocks platelet aggregation induced by ADP, collagen, thrombin, or arachidonic acid. This inhibition is found to correlate with inhibition of thromboxane A2 generation and of dense granule release of serotonin.	Applaggin activity may be assayed in vitro by measuring inhibition of platelet serotonin release induced by ADP, gamma-thrombin, and collagen. (Proc Natl Acad Sci U S A 1989 Oct;86(20):8050-4).	Thrombosis; Stroke; Ischemic Heart Disorders
Autocrine motility factor (AMF; phosphoglucose isomerase; glucose phosphate isomerase; neuroleukin)	LocusID:2821 NP_000166 XP_012854	WO8707617-A WO9909049-A1	Glucose phosphate isomerase (neuroleukin); neurotrophic factor and lymphokine; bladder cancer diagnostic.	Cell motility assay on mouse CT-26 cells (Sun et al., Proc. Natl. Acad. Sci. USA 96: 5412-5417, 1999; Lin et al., Mol. Cell. Endocrinol. 84:47-54, 1992)	Hemolytic anemia; glucosephosphate isomerase deficiency; Hydrops fetalis

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
C1 Inactivator (C1 esterase inhibitor; Berinert; Complement C1 inactivator; C1 Inattivatore Umano)	CAS-80295-38-1 LocusID:710 NP_000053 XP_006339	US5622930-A WO9106650-A	Activation of complement is an essential part of the mechanism of pathogenesis of a large number of human diseases. C1-esterase inhibitor (C1-INH) concentrate prepared from human plasma is being successfully used for the treatment of hereditary angioneurotic edema. Recently, C1-INH has been found to be consumed in severe inflammation and has been shown to exert beneficial effects in several inflammatory conditions such as human sepsis, post-operative myocardial dysfunction due to reperfusion injury, severe capillary leakage syndrome after bone marrow transplantation, reperfusion injury after lung transplantation, burn, and cytotoxicity caused by IL-2 therapy in cancer. is a major inhibitor of two pro-inflammatory plasma cascade systems, the classical pathway of complement and the contact activation system. During the activation of classical pathway, C1-INH interacts with the activated C1 and inhibits it. Interaction of C1-INH with activated C1 complex leads to the dissociation of the C1q subunit and formation	C1 inhibitor function can be assessed by measurement of inhibition of complement C1 cleavage (J Immunol 1994 Mar 15;152(6):3199-209)	Angioedema; Pancreatic Disorders; Reperfusion Injury; Transplant Rejection; Vascular Disorders

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Coagulant Complex (Anti-Inhibitor Coagulant Complex; FEIBA VH; AUTOPLEX T)					Control of spontaneous bleeding in hemophilia A and B; prevention of bleeding in patients on Factor VIII inhibitors
C-reactive protein	LocusID:1401 NP_000558 XP_001859	WO9221364-A WO9505394-A	Acute-phase serum protein that binds microbial polysaccharides and ligands on damaged cells, activates the classical complement pathway	Platelet activation (Simpson RM. Immunology 1982 47(1):193) Platelet aggregation (Cheryk LA.Vet Immunol Immunopathol 1996 52:27). Increased production of IL-1 alpha, IL-1 beta, and TNF-alpha in macrophages (Galve-de Rochemonteix B. J Leukoc Biol 1993 53:439)	
EPO (Erythropoietin; Epoetin alfa; Epoetin beta; Gene-activated erythropoietin; Darbepoetin-alpha; NESP; Epogen; Procrit; Eprex; Erypo; Espo; Epoimmun; EPOGIN; NEORECORMON; HEMOLINK; Dynepo; ARANESP)	CAS-113427-24-0 CAS-122312-54-3 LocusID:2056 NP_000790 XP_011627	WO9902710-A1 WO8502610-A WO8603520-A WO9206116-AWO9206116-A US5985607-A EP232034	Hormone that senses and regulates the level of oxygen in the blood by modulating the number of circulating erythrocytes	Cell proliferation assay using a erythroleukemic cell line TF-1. (Kitamura et al. 1989 J.Cell. Physiol. 140:323)	Anemia; Bleeding Disorders

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Factor IX (Coagulation factor IX (human); Factor IX Complex; Christmas factor; plasma thromboplastin component (PTC); prothrombin complex concentrate (PCC); Nonacog alpha; MONONINE; ALPHANINE-SD; BEBULIN; PROPLEX-T; KONYNE; PROFILNINE SD; BeneFIX; IMMUNINE VH)	CAS-181054-95-5 LocusID:2158 NP_000124 XP_010270	WO8505125-A WO8505376-A WO9747737-A1 EP162782-A WO8400560-A US4994371-A	Coagulation factor IX is a vitamin K-dependent factor that circulates in the blood as an inactive zymogen. Factor IX is converted to an active form by factor XIa, which excises the activation peptide and thus generates a heavy chain and a light chain held together by one or more disulfide bonds. In the blood coagulation cascade, activated factor IX activates factor X to its active form through interactions with Ca+2 ions, membrane phospholipids, and factor VIII. Alterations of this gene, including point mutations, insertions and deletions, cause factor IX deficiency, which is a recessive X-linked disorder, also called hemophilia B or Christmas disease.	Factor IX clotting activity: Valder R. et al., 2001 "Posttranslational modifications of recombinant myotube-synthesized human factor IX" Blood 97: 130-138.	Hemophilia B; bleeding; Factor IX deficiency; Christmas disease; bleeding episodes in patients with factor VIII inhibitor or Factor VII deficiency

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Factor VII (Coagulation Factor VII; Active-site inactivated factor VII (DEGR-VIIa/FFR-VIIa); Eptacog alfa; Coagulation Factor VIIa; Novoseven; NiaStase; Novostase; MONOCLATE-P)	CAS-102786-61-8 LocusID:2155 NP_000122 NP_062562 XP_007179 XP_007180	WO8400560-A WO9323074-A US5997864-A US5580560-A US4994371-A EP200421-A WO9427631-A WO9309804-A	Coagulation factor VII is a vitamin K-dependent factor essential for hemostasis. This factor circulates in the blood in a zymogen form, and is converted to an active form by either factor IXa, factor Xa, factor XIIa, or thrombin by minor proteolysis. Upon activation of the factor VII, a heavy chain containing a catalytic domain and a light chain containing 2 EGF-like domains are generated, and two chains are held together by a disulfide bond. In the presence of factor III and calcium ions, the activated factor then further activates the coagulation cascade by converting factor IX to factor IXa and/or factor X to factor Xa. Defects in this gene can cause coagulopathy.	Coagulation Assay using Prothrombin Clotting Time (Belaaouaj AA et al., J. Biol. Chem. 275: 27123-8, 2000; Diaz-Collier JA et al., Thromb Haemost 71: 339-46; 1994).	Bleeding Disorders; Coronary Restenosis; Hemophilia A and B; Liver Disorders; Thrombosis; Vascular Restenosis; Surgery-related hemorrhagic episodes

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Factor VIII (Factor VIII; Octocog alfa; Moroctocog alfa; Recombinant Antihemophilic factor; Nordiate; ReFacto; Kogenate; Kogenate SF; Helixate; Recombinate)	CAS-139076-62-3 LocusID:2157 NP_000123 XP_013124	WO9621035-A2 WO9703195-A1 WO9800542-A2 EP160457-A EP160457-A WO9959622-A1 EP253455-A	This gene encodes coagulation factor VIII, which participates in the intrinsic pathway of blood coagulation; factor VIII is a cofactor for factor IXa which, in the presence of Ca+2 and phospholipids, converts factor X to the activated form Xa. This gene produces two alternatively spliced transcripts. Transcript variant 1 encodes a large glycoprotein, isoform a, which circulates in plasma and associates with von Willebrand factor in a noncovalent complex. This protein undergoes multiple cleavage events. Transcript variant 2 encodes a putative small protein, isoform b, which consists primarily of the phospholipid binding domain of factor VIIIc. This binding domain is essential for coagulant activity. Defects in this gene results in hemophilia A, a common recessive X-linked coagulation disorder.	Development of a simple chromogenic factor VIII assay for clinical use. Wagenvoord RJ, Hendrix HH, Hemker HC. Haemostasis 1989;19(4):196-204.	Hemophilia A; Hemophilia; Surgery-related hemorrhagic episodes

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Factor X	LocusID:2159 NP_000495 XP_007182	WO9204378-A WO9309804-A US4994371-A	Encodes the vitamin K-dependent coagulation factor X precursor of the blood coagulation cascade. This factor precursor is converted to a mature two-chain form by the excision of the tripeptide RKR. Two chains of the factor are held together by 1 or more disulfide bonds; the light chain contains 2 EGF-like domains, while the heavy chain contains the catalytic domain which is structurally homologous to those of the other hemostatic serine proteases. The mature factor is activated by the cleavage of the activation peptide by factor IXa (in the intrinsic pathway), or by factor VIIa (in the extrinsic pathway). The activated factor then converts prothrombin to thrombin in the presence of factor Va, Ca+2, and phospholipid during blood clotting. Mutations of this gene result in factor X deficiency, a hemorrhagic condition of variable severity.	FACTOR X ACTIVITY ASSAY. Van Wijk EM et al. A rapid manual chromogenic factor X assay. Thromb Res 22, 681-686 (1981).	Factor X deficiency; Stuart-Prower factor deficiency; hemorrhage; menorrhagia; hematuria; hemarthrosis

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Factor XIII	LocusID:2162 LocusID:2163 LocusID:2164 LocusID:2165 NP_000120 NP_001985 XP_004467 XP_001350	WO9116931-A EP494702-A US7425887-A WO9102536-A WO9918200-A	Coagulation factor XIII is the last zymogen to become activated in the blood coagulation cascade. Plasma factor XIII is a heterotetramer composed of 2 A subunits and 2 B subunits. The A subunits have catalytic function, and the B subunits do not have enzymatic activity and may serve as a plasma carrier molecules. Platelet factor XIII is comprised only of 2 A subunits, which are identical to those of plasma origin. Upon activation by the cleavage of the activation peptide by thrombin and in the presence of calcium ion, the plasma factor XIII dissociates its B subunits and yields the same active enzyme, factor XIIIa, as platelet factor XIII. This enzyme acts as a transglutaminase to catalyze the formation of gamma-glutamyl-epsilon-lysine crosslinking between fibrin molecules, thus stabilizing the fibrin clot. Factor XIII deficiency is classified into two categories: type I deficiency, characterized by the lack of both the A and B subunits; and type II deficiency, characterized by the lack of the A subunit alone	BLOOD COAGULATION ASSAY. Karpati L, Penke B, Katona E, Balogh I, Vamosi G, Muszbek L. A modified, optimized kinetic photometric assay for the determination of blood coagulation factor XIII activity in plasma. Clin Chem. 2000 Dec;46(12):1946-55.	Factor XIII deficiency; bleeding tendency; defective wound healing; habitual abortion

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Fibrinogen; thrombin; aprotinin (Human fibrinogen; human thrombin; aprotinin; and calcium chloride; synthocytes; FAMs; BERIPLAST-P)	LocusID:2243 LocusID:2244 LocusID:2266 LocusID:2147 NP_000499 NP_068657 NP_005132 NP_000500 NP_068656 NP_000497	US6083902-A WO9523868-A1 WO9416085-A WO9523868-A1 WO9523868-A1 WO9529686-A1 US6083902-A WO9523868-A1 WO9528946-A1 WO9313208-A US5502034-A US5476777-A US6110721-A	Following vascular injury, fibrinogen is cleaved by thrombin to form fibrin, which is the most abundant component of blood clots. In addition, various cleavage products of fibrinogen and fibrin regulate cell adhesion and spreading, display vasoconstrictor and chemotactic activities, and are mitogens for several cell types. Coagulation factor II is proteolytically cleaved to form thrombin in the first step of the coagulation cascade, which ultimately results in the stemming of blood loss. F2 also plays a role in maintaining vascular integrity.	Fibrinogen assay: Halbmayer WM, Haushofer A, Schon R, Radek J, Fischer M. Comparison of a new automated kinetically determined fibrinogen assay with the 3 most used fibrinogen assays (functional, derived and nephelometric) in Austrian laboratories in several clinical populations and healthy controls. Haemostasis 1995 May-Jun;25(3):114-23. Tan V, Doyle CJ, Budzynski AZ. Comparison of the kinetic fibrinogen assay with the von Clauss method and the clot recovery method in plasma of patients with conditions affecting fibrinogen coagulability. Am J Clin Pathol. 1995 Oct;104(4):455-62. Lawrie AS, McDonald SJ, Purdy G, Mackie IJ, Machin SJ. Prothrombin time derived fibrinogen determination on Sysmex CA-6000. J Clin Pathol. 1998 Jun;51(6):462-6. Aprotinin assay: Cardigan RA, Mackie IJ, Gippner-Steppert C, Jochum M, Royston D, Gallimore MJ. Determination of plasma aprotinin levels by functional and immunologic assays. Blood Coagul Fibrinolysis. 2001 Jan;12(1):37-42. Thrombin assay: Syed S, R[4]C PD, Kulczycky M, Sheffield WP. Potent antithrombin activity and delayed clearance from the circulation characterize recombinant hirudin genetically fused to albumin. Blood. 1997 May 1;89(9):3243-52.	Tissue adhesion; thrombosis; bleeding disorders; wounds; thrombocytopenia; dysfibrinogenemia; hypofibrinogenemia; afibrinogenemia; renal amyloidosis; thrombosis; dysprothrombinemia

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
G-CSF (Granulocyte colony-stimulating factor; Granulokine; KRN 8601; Filgrastim; Lenograstim; Meograstim; Nartograstim; Neupogen; NOPIA; Gran; GRANOCYTE; Granulokine; Neutrogin; Neu-up; Neutromax)	CAS-121181-53-1 CAS-135968-09-1 CAS-130120-55-7 CAS-130120-54-6 CAS-134088-74-7 LocusID:1440 NP_000750 XP_008227	WO8604506-A EP220520-A WO8604506-A WO8701132-A US6054294-A	Stimulates the proliferation and differentiation of the progenitor cells for granulocytes	Proliferation of murine NFS-60 cells (Weinstein et al, Proc Natl Acad Sci U S A 1986; 83, pp5010-4)	Chemoprotection; Inflammatory disorders; Myelocytic leukemia; Primary neutropenias (e.g.; Kostmann syndrome) or secondary neutropenia; Prevention of neutropenia; Prevention and treatment of neutropenia in HIV-infected patients; Infections associated with neutropenias; Myelopysplasia; Autoimmune disorders
GM-CSF (Granulocyte-macrophage colony-stimulating factor; rhuGM-CSF; BI 61012; Prokine; Molgramostim; Sargramostim; GM-CSF/IL 3 fusion; Milodistim; Leucotropin; PROKINE; LEUKOMAX; Interberin; Leukine; Leukine Liquid; Pixykine)	CAS-99283-10-0 CAS-123774-72-1 CAS-60154-12-3 CAS-137463-76-4 LocusID:1437 NP_000749 XP_003751	WO8805786 WO8600639 WO8603225 US5391706 US5545536	Regulates hematopoietic cell differentiation, gene expression, growth	Colony Stimulating Assay: Testa, N.G., et al., "Assays for hematopoietic growth factors." Balkwill FR (edt) Cytokines, A practical Approach, pp 229-44; IRL Press Oxford 1991.	Bone Marrow Disorders; Bone marrow transplant; Chemoprotection; Hepatitis C; HIV Infections; Lung Cancer; Malignant melanoma; Mycobacterium avium complex; Mycoses; Myeloid Leukemia; Neonatal infections; Neutropenia; Oral mucositis; Prostate Cancer; Stem Cell Mobilization; Vaccine Adjuvant; Venous Stasis Ulcers; Prevention of neutropenia; Acute myelogenous leukemia; Hematopoietic progenitor cell mobilization; Non-Hodgkin's lymphoma; Acute lymphoblastic leukemia; Hodgkin's disease; Accelerated myeloid recovery; Xenotransplant rejection

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Hepatitis IG (Hepatitis B Immune Globulin; Hepatitis C immune globulin; HCVIG; BAYHEP; NABI-HB; NABI-CIVACIR)					Exposure to Hepatitis B (HBsAg) or Hepatitis C; perinatal exposure of infants with HBV or HCV infected mothers;; sexual or household exposure to patient with acute HBV or HCV

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Hepatocyte growth factor (HGF; Scatter factor; SF; HGF/SF)	LocusID:3082 Genbank:CAA34387	JP03130091-A JP10070990-A WO9323541-A	HGF disrupts desmosomal junctions between epithelial cells and induces a motile fibroblast-like phenotype in individual cells. The factor therefore also influences the invasive growth of tumor cells derived from epithelial cells and may be involved also in processes of Wound healing and early embryonic development. For some cell types including keratinocytes and mammary epithelial cells HGF is merely a motility factor. It is also an autocrine modulator that influences the motility of the cells that produce it. It is a potent mitogen for hepatocytes and also a morphogen (see: HGF , hepatocyte growth factor). HGF binds to heparin and this may be important for its activities in vivo. The actions of HGF are inhibited by Suramin . HGF has been shown to be an Angiogenesis factor in vivo. It induces cultured microvascular endothelial cells to accumulate and secrete significantly increased quantities of urokinase, an enzyme associated with development of an invasive endothelial phenotype during angiogenesis . HGF	Adams JC et al Production of scatter factor by ndk, a strain of epithelial cells, and inhibition of scatter factor activity by suramin. Journal of Cell Science 98: 385-94 (1991); Bhargava MM et al Purification, characterization and mechanism of action of scatter factor from human placenta. Experientia Suppl. 59: 63-75 (1991); Coffer A et al Purification and characterization of biologically active scatter factor from ras-transformed NIH-3T3 conditioned medium. Biochemical Journal 278: 35-41 (1991); Dowrick PG et al Scatter factor affects major changes in the cytoskeletal organization of epithelial cells Cytokine 3: 299-310 (1991); Furlong RA et al Comparison of biological and immunochemical properties indicates that scatter factor and hepatocyte growth factor are indistinguishable. Journal of Cell Science 100: 173-7 (1991); Gherardi E et al Purification of scatter factor, a fibroblast-derived basic protein that modulates epithelial interactions and movement. Proceedings of the National Academy of Science (USA)	Alopecia; Cancer; Chemoprotection; Cirrhosis; Haematological disorders; Radioprotection

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Hirudin (Lepirudin; Desirudin; Refludan; Revasc)	CAS-138068-37-8 CAS-120993-53-5 CAS-8001-27-2 Genbank:AAA29195 Genbank:AAA01384	WO8504418-A EP200655-A EP503829-A WO9207874-A WO9201712-A EP340170-A EP341215-A WO9207874-A WO9201712-A	Hirudin is a potent anticoagulant which inhibits thrombin.	Hirudin activity can be measured using a platelet aggregation assay (Blood Coagul Fibrinolysis 1996 Mar;7(2):259-61).	Coronary restenosis; Deep Vein Thrombosis; Disseminated Intravascular Coagulation; Heparin-induced thrombocytopenia and thrombosis syndrome; Myocardial infarction; Unstable Angina Pectoris; Anticoagulant in adults suffering from acute coronary syndrome; Thrombosis; Veinous Thrombosis

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Human growth hormone (Pegvisamont; Somatrem; Somatropin; TROVERT; PROTROPIN; BIO-TROPIN; HUMATROPE; NUTROPIN; NUTROPIN AQ; NUTROPHIN; NORDITROPIN; GENOTROPIN; SAIZEN; SEROSTIM)	CAS-82030-87-3 CAS-12629-01-5 LocusID:2688 LocusID:2689 NP_000506 NP_072053 NP_072054 NP_072055 NP_072056 NP_002050 NP_072050 NP_072051 NP_072052 XP_008250	WO9418227-A WO9005185-A WO9520398-A EP245138-A WO8605804-A WO9004788-A WO9418227-A1 US6013579-A US6194176-A WO8605804 US6110707 US4977089 US5580723 US5955346 US6013478 WO8605804 WO9004788-A WO9418227-A1 US6110707-A US4977089 US5580723 US5955346 US6013478 WO8605804 US6110707 US5580723 US5955346 US6013478 WO8605804 WO9004788-A WO9418227-A1 US6110707-A US4977089 US5580723 US5955346 US6013478	Plays an important role in growth control; binds 2 GHR molecules and induces signal transduction through receptor dimerization	Ba/F3-hGHR proliferation assay, a novel specific bioassay for serum human growth hormone. J Clin Endocrinol Metab 2000 Nov;85(11):4274-9 Plasma growth hormone (GH) immunoassay and tibial bioassay, Appl Physiol 2000 Dec;89(6):2174-8 Growth hormone (hGH) receptor mediated cell mediated proliferation, Growth Horm IGF Res 2000 Oct;10(5):248-55 International standard for growth hormone, Horm Res 1999;51 Suppl 1:7-12	Acromegaly; Growth failure; Growth failure and endogenous growth hormone replacement; Growth hormone deficiency; Growth failure and growth retardation Prader-Willi syndrome in children 2 years or older; Growth deficiencies; Postmenopausal osteoporosis; burns; cachexia; cancer cachexia; dwarfism; metabolic disorders; obesity; renal failure; Turner's Syndrome; fibromyalgia; fracture treatment; frailty

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Insulin (Human insulin; Insulin aspart; Insulin Glargine; Insulin lispro; Lys-B28 Pro-B29; lyspro; LY 275585; diarginylinsulin; Des-B26-B30-insulin-B25-amide; Insulin detemir; LABI; NOVOLIN; NOVORAPID; HUMULIN; NOVOMIX 30; VELOSULIN; NOVOLOG; LANTUS; ILETIN; HUMALOG; MACRULIN; EXUBRA; INSUMAN; ORALIN; ORALGEN; HUMAHALE; HUMAHALIN)	CAS-11061-68-0 CAS-116094-23-6 CAS-133107-64-9 CAS-160337-95-1 LocusID:3630 NP_000198 XP_006400	WO200040613-A1 EP37723-A EP55942-A US4431740-A US4430266-A US4624926-A US5077204-A US5840542-A US6110707-A WO9200322-A	Insulin is a heterodimeric polypeptide hormone involved in carbohydrate metabolism. After removal of the precursor signal peptide, proinsulin is post-translationally cleaved into two chains (peptide A and peptide B) that are covalently linked via two disulfide bonds. Binding of this mature form of insulin to the insulin receptor (INSR) stimulates glucose uptake.	Insulin activity may be assayed in vitro using a [3-H]-glucose uptake assay. (J Biol Chem 1999 Oct 22; 274(43):30864-30873).	Hyperglycemia; Diabetes mellitus; Type 1 diabetes and type 2 diabetes

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Interferon alfa (Interferon alfa-2b; recombinant; Interferon alfa-n1; Interferon alfa-n3; Peginterferon alpha-2b; Ribavirin and interferon alfa-2b; Interferon alfacon-1; interferon consensus; YM 643; CIFN; interferon -alpha consensus; recombinant methionyl consensus interferon; recombinant consensus interferon; CGP 35269; RO 253036; RO 258310; INTRON A; PEG-INTRON; OIF; OMNIFERON; PEG-OMNIFERON; VELDONA; PEG-REBETRON; ROFERON A; WELLFERON; ALFERON N/LDO; REBETRON; ALTEMOL; VIRAFERONPEG; PEGASYS; VIRAFERON; VIRAFON; AMPLIGEN; INFERGEN; INFAREX; ORAGEN)	CAS-74899-72-2 CAS-76543-88-9 CAS-99210-65-8 LocusID:3440 NP_000596 XP_011801 CAS-9008-11-1 CAS-118390-30-0 LocusID:3439 NP_076918	EP32134 WO9419373-A WO9201055 US5602232 US6069133 WO8302461 US6069133 US4569908 US4758428	Interferon alpha belongs to the type I Interferon family of functionally related cytokines that confer a range of cellular responses including antiviral, antiproliferative, antitumor and immunomodulatory activities.	Anti-viral assay: Rubinstein S, Familletti PC, Pestka S. (1981) Convenient assay for interferons. J. Virol. 37(2):755-8; Anti-proliferation assay: Gao Y, et al (1999) Sensitivity of an epstein-barr virus-positive tumor line, Daudi, to alpha interferon correlates with expression of a GC-rich viral transcript. Mol Cell Biol. 19(11):7305-13.	Hepatitis C; oncology uses; cancer; hepatitis; human papilloma virus; fibromyalgia; Sjogren's syndrome; hairy cell leukemia; chronic myelogeonus leukemia; AIDS-related Kaposi's sarcoma; chronic hepatitis B; malignant melanoma; non-Hodgkin's lymphoma; external condylomata acuminata; HIV infection; small cell lung cancer; hematological malignancies; herpes simplex virus infections; multiple sclerosis; viral hemmorhagic fevers; solid tumors; renal cancer; bone marrow disorders; bone disorders; bladder cancer; gastric cancer; Hepatitis D; multiple myeloma; type I diabetes mellitus; viral infections; Cutaneous T-cell lymphoma; Cervical dysplasia; Chronic fatigue syndrome; Renal cancer

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
IVIG (Intravenous Immune Globulin; VENOGLOBULIN-S; PANGLOBULIN; POLYGAM; GAMMAR-P; GAMMAGARD S/D; IVEEGAM; BAYGAN; SANDOGLOBULIN; GAMIMUNE)			Regulates hematopoietic cell differentiation, gene expression, growth		Immune deficiencies; agammaglobulinemia; hypogammaglobulinemia; immunodeficient states and bacterial infections; Kawasaki Syndrome; Hepatitis A; measles varicella; rubella; immunoglobulin deficiency; idiopathic thrombocytopenic purpura; primary humoral immunodeficiency states; bone marrow transplantation; pediatric HIV infection; Guillain-Barre syndrome; chronic inflammatory demyelinating polyneuropathy; multifocal neuropathy; dermatomyositis; amyotrophic lateral sclerosis; inclusion-body myositis; Lambert-Eaton myasthenic syndrome; Rasmussen syndrome; West syndrome; intractable childhood epilepsy; Lennox-Gastaut syndrome; polymyositis; relapsing-remitting multiple sclerosis; optic neuritis; stiff-man syndrome; paraneuplastic cerebellar degeneration; paraneoplastic encephalomyelitis and sensory neuropathy; systemic vasculitis; myelopathy associated with human T-cell lymphotrophic virus-1 infection.

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
IVIG-CMV (Cytomegalovirus immune globulin intravenous (human); CMV_IVIG; CYTOGAM)					Cytomegalovirus disease
Laminin	LocusID:3907 LocusID:3908 LocusID:3909 LocusID:3910 LocusID:3911 LocusID:3912 LocusID:3913 LocusID:3914 LocusID:3915 LocusID:3918 LocusID:10319 NP_000417 NP_000218 NP_002281 NP_002282 NP_002283 NP_000219 NP_002284 NP_005553 NP_061486 NP_006050 XP_011387 XP_008772 XP_004301 XP_011616 XP_001716 XP_002204 XP_002202 XP_002203 XP_011791	WO9506660 US5658789-A WO8901493-A WO9811217-A2 WO9511972-A WO9111462-A WO9508628-A2 WO200066732-A2 WO9815179-A1 WO9919348-A1 WO200066730-A2 US7267564-A WO9610646-A1 WO200066731-A2 US5658789-A WO200058473-A2	Basement membrane protein; cell adhesion, differentiation, migration, signaling, neurite outgrowth and metastasis	Neurite outgrowth assay, Neurosci Lett 2001 Mar 30;301(2):83-6 Cell adhesion assay, CAFCA, Centrifugal Assay for Fluorescence-based Cell Adhesion, Cancer Res 2001 Jan 1;61(1):339-47 Cell migration assay, Biochem Biophys Res Commun 2000 Nov 30;278(3):614-20	

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
MCP/MCAF (Monocyte Chemotactic Protein; Monocyte chemoattracting peptides)	LocusID:6347 LocusID:6355 LocusID:6354 NP_002973 NP_005614 NP_006264 XP_008415 XP_008412 XP_012649	US5714578-A US7330446-A US6090795-A US7304234-A US5571713-A US5605671-A WO9725427-A1 EP906954-A1 WO9912968-A2 WO9509232-A EP488900-A WO9504158-A WO9509232-A	Chemotactic factor for monocytes; chemokine involved in recruiting leukocytes during inflammation; attracts macrophages during inflammation and metastasis	Transendothelial lymphocyte chemotaxis assay: Carr, M.W., et al., Proc. Natl. Acad. Sci. USA, vol. 91, pp. 3652-3656 (April 1994).	Cancer; Chemoprotection; Wounds

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
M-CSF (Macrophage colony-stimulating factor; CSF-1; Cilmostim; Macstim)	CAS-148637-05-2 LocusID:1435 NP_000748 XP_002150	US5171675-A US4929700-A US5573930-A US5672343-A US5681719-A US5643563-A US5861150-A US6117422-A US6103224-A US6156300-A US6146851-A	M-CSF stimulates the growth of macrophage/granulocyte-containing colonies in soft agar cultures, influences the proliferation and differentiation of hematopoietic stem cells into macrophages but mainly the growth survival and differentiation of monocytes. In combination with another colony stimulating factor, GM-CSF, one observes the phenomenon of synergistic suppression, i. e., the combination of these two factors leads to a partial suppression of the generation of macrophage-containing cell colonies. M-CSF is a specific factor in that the proliferation inducing activity is more or less restricted to the macrophage lineage. M-CSF also is a potent stimulator of functional activities of monocytes. In normal human macrophages M-CSF induces antibody-dependent cellular cytotoxicity. In monocytes and macrophages M-CSF induces the synthesis of IL1, G-CSF, IFN, TNF, plasminogen activator, thromboplastin, prostaglandins and thromboxanes.	M-CSF can be assayed in a Colony formation assay by the development of colonies containing macrophages (Int J Cell Cloning 1984 Nov;2(6):356-67). M-CSF is also detected in specific Bioassays with cells lines that depend in their growth on the presence of M-CSF or that respond to this factor, for example, BAC1.2F5 ; BaF3 ; GNFS-60 ; J774. An alternative and entirely different detection method is RT-PCR quantitation of cytokines.	Cancer; Hypercholesterolemia
MSF (Migration stimulating factor)	Genbank:CAC20427	WO9931233-A1	Chemotaxis	Transendothelial lymphocyte chemotaxis assay: Carr, M.W., et al., Proc. Natl. Acad. Sci. USA, vol. 91, pp. 3652-3656 (April 1994).	Wound healing

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
NCAF (Neutrophil chemoattracting peptides)					
Osteopontin (OPN, BNSP, BSPI, ETA-1,secreted phosphoprotein 1, bone sialoprotein I, early T-lymphocyte activation 1)	LocusID:6696 NP_000573 XP_011125	WO9915904-A1 WO200062065-A1 WO9222316-A	Osteopontin (OPN) is a highly phosphorylated sialoprotein that is a prominent component of the mineralized extracellular matrices of bones and teeth. OPN is characterized by the presence of a polyaspartic acid sequence and sites of Ser/Thr phosphorylation that mediate hydroxyapatite binding, and a highly conserved RGD motif that mediates cell attachment/signaling. Expression of OPN in a variety of tissues indicates a multiplicity of functions that involve one or more of these conserved motifs. OPN is involved in a range of biological activities including developmental processes, wound healing, immunological responses, tumorigenesis, bone resorption, and calcification.	Cell Attachment Assay: Senger DR, Perruzzi CA, Papadopoulos-Sergiou A, Van de Water L. (1994) Adhesive properties of osteopontin: regulation by a naturally occurring thrombin-cleavage in close proximity to the GRGDS cell-binding domain. Mol Biol Cell 5(5):565-74	Bone Fractures
Platelet Factor 4 (Endostatin B; Iroplact; RG 1001; Replistatin)	CAS-37270-94-3 LocusID:5196 NP_002610 XP_003505	WO9302192-A WO9504158-A US5248666-A US5776892-A	Platelet factor 4 (PF-4) is a CXC-chemokine with strong anti-angiogenic properties.	Anti-angeogenic assay (PMID: 11259363)	Bleeding Disorders; Colorectal Cancer; Diabetic Retinopathy; Glioma; Heparin Neutralization after Cardiac Catheterization or Cardiopulmonary Bypass Surgery; Kaposi's Sarcoma; Malignant Melanoma; Renal Cancer

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Protein C (Drotrecogin alfa; Activated Protein C; CTC 111; Ceprotin; rhAPC; Zovant)	CAS-60202-16-6 LocusID:5624 NP_000303 XP_002706	WO9109953-A WO9112320-A US5516650-A US5358932-A	Protein C is a serine protease involved in coagulation and fibrinolysis.	Protein C activity may be assayed in vitro using a coagulation assay. (J Biol Chem 2000 Sep 1; 275(35): 27123-27128; Thromb Haemost 1994 Mar;71(3):339-346).	Disseminated intravascular coagulation; Septic shock; Thrombosis
Prothrombin (Factor II, Thrombin, F2)	LocusID:2147 NP_000497	WO9313208-A US5502034-A US5476777-A US6110721-A	Coagulation factor II is proteolytically cleaved to form thrombin in the first step of the coagulation cascade, which ultimately results in the stemming of blood loss. F2 also plays a role in maintaining vascular integrity.	Prothrombin quantitation and activation assay. "CA-1 method, a novel assay for quantification of normal prothrombin using a Ca2+ - dependent prothrombin activator, carinactivase-1." Thromb Res. 1999 May 15;94(4):221-6. "Activation of human prothrombin by arginine-specific cysteine proteinases (Gingipains R) from Porphyromonas gingivalis." J Biol Chem. 2001 Mar 16	
Rabies IG (Rabies Immune Globulin; BAYRAB; HYPERRAB; IMOGAM RABIES-HT; IMOGAM)					Rabies
RhoD IG (RhoD Immune Globulin; IVIG-Rho(D); PAYRHO-D; MICRHOGAM; RHOGAM; WinRho SDF)					Prevention of isoimmunization of RhoD negative women at time of spontaneous or induced abortion or transfusion in pregnancy; Hemolyic disease; immune thrombocytopenic purpura; HIV infection
RSV IVIG (Respiratory syncytial virus IV immune globulin (human); Hypermune RSV; RESPIGAM)					Lower respiratory tract infections; respiratory syncytial infections

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Serum Cholinesterase	LocusID:590 LocusID:1110 NP_000046 XP_003134	WO9107483-A WO9523158-A US6001625-A US5695750-A	Also known as butyrylcholinesterase/pseuocholinesterase E1(CHE1). Human tissues have two distinct cholinesterase activities: acetylcholinesterase and butyrylcholinesterase. Acetylcholinesterase functions in the transmission of nerve impulses, whereas the physiological function of butyryl-cholinesterase remains unknown. An atypical form of butyrylcholinesterase or the absence of its activity leads to prolonged apnea following administration of the muscle relaxant suxamethonium. The widespread expression of CHE1 in early differentiation suggests development-related functions for this protein.	BchE activity assay "Differential inhibition of human serum cholinesterase with fluoride: recognition of two new phenotypes." Nature 191: 496-498, 1961. "A rare genetically determined variant of pseudocholinesterase in two German families with high plasma enzyme activity." Europ. J. Biochem. 99: 65-69, 1979. "Genetic analysis of a Japanese patient with butyrylcholinesterase deficiency." Ann. Hum. Genet. 61:491-496, 1997.	
Tenascin	LocusID:7143 LocusID:7146 LocusID:7148 NP_003276 NP_009047 XP_001730 XP_004201	WO9628550-A1 WO9608513-A1 US5681931-A US5635360-A WO9222319-A WO9608513-A1 US5681931-A US5635360-A US6048704-A	The tenascins (TN) are a family of extracellular matrix proteins. The genes are expressed in distinct tissues at different times during embryonic development and are present in adult tissues. TN-R is detected predominantly in the central nervous system of early embryos and likely involved in central nervous system development. TN-XA is overexpressed in many tumors.	cell adhesion assay. "Cell adhesion to fibronectin and tenascin: quantitative measurements of initial binding and subsequent strengthening response."J Cell Biol. 1989 Oct;109(4 Pt 1):1795-805.; "Tenascin interferes with fibronectin action." Cell. 1988 May 6;53(3):383-90. neurite growth in vitro. "Tenascin is accumulated along developing peripheral nerves and allows neurite outgrowth in vitro." Development. 1990 Oct;110(2):401-15.	

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Tetanus IG (Tetanus Immune Globulin; TIG; BAYTET)					Tetanus
Vitronectin	LocusID:7448 NP_000629 XP_008484	US5514582-A US6140072-A WO9213075-A	Binds to serpin serine protease inhibitors such as PAI, mediates cell-to-substrate adhesion, inhibits the cytolytic action of the terminal complement cascade in vitro and inhibits inactivation of thrombin by antithrombin, thereby regulating coagulation.	Individual functions of the molecule are assayed separately. The cell adhesion function is assayed using a cell adhesion assay (Feinberg and Vogelstein,1983; Anal. Biochem. 132 pp6-10); PAI binding using a solid phase binding assay (Seiffert & Loskutoff, 1991; J. Biol. Chem. 266 pp2824-2830)	Atherosclerosis; Vascular Restenosis; Cancer; Cardiovascular Disorders; Malignant Melanoma; Clotting disorders; Transplantation.

In preferred embodiments, the albumin fusion proteins of the invention are capable of a therapeutic activity and/or biologic activity corresponding to the therapeutic activity and/or biologic activity of the Therapeutic protein corresponding to the Therapeutic protein portion of the albumin fusion protein listed in the corresponding row of Table 1. (See, e.g., the "Biological Activity" and "Therapeutic Protein X"columns of Table 1.) In further preferred embodiments, the therapeutically active protein portions of the albumin fusion proteins of the invention are fragments or variants of the reference sequence cited in the "Exemplary Identifier" column of Table 1, and are capable of the therapeutic activity and/or biologic activity of the corresponding Therapeutic protein disclosed in "Biological Activity" column of Table 1.

Polypeptide and Polynucleotide Fragments and Variants

Fragments

The present invention is further directed to fragments of the Therapeutic proteins described in Table 1, albumin proteins, and/or albumin fusion proteins of the invention.

Even if deletion of one or more amino acids from the N-terminus of a protein results in modification or loss of one or more biological functions of the Therapeutic protein, albumin protein, and/or albumin fusion protein, other Therapeutic activities and/or functional activities (e.g., biological activities, ability to multimerize, ability to bind a ligand) may still be retained. For example, the ability of polypeptides with N-terminal deletions to induce and/or bind to antibodies which recognize the complete or mature forms of the polypeptides generally will be retained when less than the majority of the residues of the complete polypeptide are removed from the N-terminus. Whether a particular polypeptide lacking N-terminal residues of a complete polypeptide retains such immunologic activities can readily be determined by routine methods described herein and otherwise known in the art. It is not unlikely that a mutein with a large number of deleted N-terminal amino acid residues may retain some biological or immunogenic activities. In fact, peptides composed of as few as six amino acid residues may often evoke an immune response.

Accordingly, fragments of a Therapeutic protein corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention, include the full length protein as well as polypeptides having one or more residues deleted from the amino terminus of the amino acid sequence of the reference polypeptide (e.g., a Therapeutic protein as disclosed in Table 1). In particular, N-terminal deletions may be described by the general formula m-q, where q is a whole integer representing the total number of amino acid residues in a reference polypeptide (e.g., a Therapeutic protein referred to in Table 1), and m is defined as any integer ranging from 2 to q-6. Polynucleotides encoding these polypeptides are also

encompassed by the invention.

In addition, fragments of serum albumin polypeptides corresponding to an albumin protein portion of an albumin fusion protein of the invention, include the full length protein as well as polypeptides having one or more residues deleted from the amino terminus of the amino acid sequence of the reference polypeptide (i.e., serum albumin). In particular, N-terminal deletions may be described by the general formula m-585, where 585 is a whole integer representing the total number of amino acid residues in serum albumin (SEQ ID NO:18), and m is defined as any integer ranging from 2 to 579. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Moreover, fragments of albumin fusion proteins of the invention, include the full length albumin fusion protein as well as polypeptides having one or more residues deleted from the amino terminus of the albumin fusion protein. In particular, N-terminal deletions may be described by the general formula m-q, where q is a whole integer representing the total number of amino acid residues in the albumin fusion protein, and m is defined as any integer ranging from 2 to q-6. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Also as mentioned above, even if deletion of one or more amino acids from the N-terminus or C-terminus of a reference polypeptide (e.g., a Therapeutic protein and/or serum albumin protein) results in modification or loss of one or more biological functions of the protein, other functional activities (e.g., biological activities, ability to multimerize, ability to bind a ligand) and/or Therapeutic activities may still be retained. For example the ability of polypeptides with C-terminal deletions to induce and/or bind to antibodies which recognize the complete or mature forms of the polypeptide generally will be retained when less than the majority of the residues of the complete or mature polypeptide are removed from the C-terminus. Whether a particular polypeptide lacking the N-terminal and/or C-terminal residues of a reference polypeptide retains Therapeutic activity can readily be determined by routine methods described herein and/or otherwise known in the art.

The present invention further provides polypeptides having one or more residues deleted from the carboxy terminus of the amino acid sequence of a Therapeutic protein corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention (e.g., a Therapeutic protein referred to in Table 1). In particular, C-terminal deletions may be described by the general formula 1-n, where n is any whole integer ranging from 6 to q-1, and where q is a whole integer representing the total number of amino acid residues in a reference polypeptide (e.g., a Therapeutic protein referred to in Table 1). Polynucleotides encoding these polypeptides are also encompassed by the invention.

In addition, the present invention provides polypeptides having one or more residues deleted from the carboxy terminus of the amino acid sequence of an albumin protein

corresponding to an albumin protein portion of an albumin fusion protein of the invention (e.g., serum albumin). In particular, C-terminal deletions may be described by the general formula 1-n, where n is any whole integer ranging from 6 to 584, where 584 is the whole integer representing the total number of amino acid residues in serum albumin (SEQ ID NO:18) minus 1. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Moreover, the present invention provides polypeptides having one or more residues deleted from the carboxy terminus of an albumin fusion protein of the invention. In particular, C-terminal deletions may be described by the general formula 1-n, where n is any whole integer ranging from 6 to q-1, and where q is a whole integer representing the total number of amino acid residues in an albumin fusion protein of the invention. Polynucleotides encoding these polypeptides are also encompassed by the invention.

In addition, any of the above described N- or C-terminal deletions can be combined to produce a N- and C-terminal deleted reference polypeptide. The invention also provides polypeptides having one or more amino acids deleted from both the amino and the carboxyl termini, which may be described generally as having residues m-n of a reference polypeptide (e.g., a Therapeutic protein referred to in Table 1, or serum albumin (e.g., SEQ ID NO:18), or an albumin fusion protein of the invention) where n and m are integers as described above. Polynucleotides encoding these polypeptides are also encompassed by the invention.

The present application is also directed to proteins containing polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to a reference polypeptide sequence (e.g., a Therapeutic protein, serum albumin protein or an albumin fusion protein of the invention) set forth herein, or fragments thereof. In preferred embodiments, the application is directed to proteins comprising polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to reference polypeptides having the amino acid sequence of N- and C-terminal deletions as described above. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Preferred polypeptide fragments of the invention are fragments comprising, or alternatively, consisting of, an amino acid sequence that displays a Therapeutic activity and/or functional activity (e.g. biological activity) of the polypeptide sequence of the Therapeutic protein or serum albumin protein of which the amino acid sequence is a fragment.
Other preferred polypeptide fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Variants

"Variant" refers to a polynucleotide or nucleic acid differing from a reference nucleic acid or polypeptide, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the reference nucleic acid or polypeptide.

As used herein, "variant", refers to a Therapeutic protein portion of an albumin fusion protein of the invention, albumin portion of an albumin fusion protein of the invention, or albumin fusion protein differing in sequence from a Therapeutic protein (e.g. see "therapeutic" column of Table 1), albumin protein, and/or albumin fusion protein of the invention, respectively, but retaining at least one functional and/or therapeutic property thereof (e.g., a therapeutic activity and/or biological activity as disclosed in the "Biological Activity" column of Table 1) as described elsewhere herein or otherwise known in the art. Generally, variants are overall very similar, and, in many regions, identical to the amino acid sequence of the Therapeutic protein corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention, albumin protein corresponding to an albumin protein portion of an albumin fusion protein of the invention, and/or albumin fusion protein of the invention. Nucleic acids encoding these variants are also encompassed by the invention.

The present invention is also directed to proteins which comprise, or alternatively consist of, an amino acid sequence which is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100%, identical to, for example, the amino acid sequence of a Therapeutic protein corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention (e.g., an amino acid sequence disclosed in the "Exemplary Identifier" column of Table 1, or fragments or variants thereof), albumin proteins (e.g., SEQ ID NO:18 or fragments or variants thereof) corresponding to an albumin protein portion of an albumin fusion protein of the invention, and/or albumin fusion proteins of the invention. Fragments of these polypeptides are also provided (e.g., those fragments described herein). Further polypeptides encompassed by the invention are polypeptides encoded by polynucleotides which hybridize to the complement of a nucleic acid molecule encoding an amino acid sequence of the invention under stringent hybridization conditions (e.g., hybridization to filter bound DNA in 6X Sodium chloride/Sodium citrate (SSC) at about 45 degrees Celsius, followed by one or more washes in 0.2X SSC, 0.1% SDS at about 50 - 65 degrees Celsius), under highly stringent conditions (e.g., hybridization to filter bound DNA in 6X sodium chloride/Sodium citrate (SSC) at about 45 degrees Celsius, followed by one or more washes in 0.1X SSC, 0.2% SDS at about 68 degrees Celsius), or under other stringent hybridization conditions which are known to those of skill in the art (see, for example, Ausubel, F.M. et al., eds:, 1989 *Current protocol in Molecular Biology*, Green publishing associates, Inc., and John Wiley & Sons Inc., New York, at pages 6.3.1 - 6.3.6 and 2.10.3). Polynucleotides encoding these polypeptides are also encompassed by the invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical"

to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, or substituted with another amino acid. These alterations of the reference sequence may occur at the amino- or carboxy-terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequence of an albumin fusion protein of the invention or a fragment thereof (such as the Therapeutic protein portion of the albumin fusion protein or the albumin portion of the albumin fusion protein), can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci.6:237-245 (1990)). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is expressed as percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues

to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C- terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

The variant will usually have at least 75 % (preferably at least about 80%, 90%, 95% or 99%) sequence identity with a length of normal HA or Therapeutic protein which is the same length as the variant. Homology or identity at the nucleotide or amino acid sequence level is determined by BLAST (Basic Local Alignment Search Tool) analysis using the algorithm employed by the programs blastp, blastn, blastx, tblastn and tblastx (Karlin *et al.*, Proc. Natl. Acad. Sci. USA 87: 2264-2268 (1990) and Altschul, J. Mol. Evol. 36: 290-300 (1993), fully incorporated by reference) which are tailored for sequence similarity searching.

The approach used by the BLAST program is to first consider similar segments between a query sequence and a database sequence, then to evaluate the statistical significance of all matches that are identified and finally to summarize only those matches which satisfy a preselected threshold of significance. For a discussion of basic issues in similarity searching of sequence databases, see Altschul *et al.*, (Nature Genetics 6: 119-129 (1994)) which is fully incorporated by reference. The search parameters for histogram, descriptions, alignments, expect (i.e., the statistical significance threshold for reporting matches against database sequences), cutoff, matrix and filter are at the default settings. The default scoring matrix used by blastp, blastx, tblastn, and tblastx is the BLOSUM62 matrix (Henikoff *et al.*, Proc. Natl. Acad. Sci. USA 89: 10915-10919 (1992), fully incorporated by reference). For blastn, the scoring matrix is set by the ratios of M (i.e., the reward score for a pair of matching residues)

to N (i.e., the penalty score for mismatching residues), wherein the default values for M and N are 5 and -4, respectively. Four blastn parameters may be adjusted as follows: Q=10 (gap creation penalty); R=10 (gap extension penalty); wink=1 (generates word hits at every winkth position along the query); and gapw=16 (sets the window width within which gapped alignments are generated). The equivalent Blastp parameter settings were Q=9; R=2; wink=1; and gapw=32. A Bestfit comparison between sequences, available in the GCG package version 10.0, uses DNA parameters GAP=50 (gap creation penalty) and LEN=3 (gap extension penalty) and the equivalent settings in protein comparisons are GAP=8 and LEN=2.

The polynucleotide variants of the invention may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, polypeptide variants in which less than 50, less than 40, less than 30, less than 20, less than 10, or 5-50, 5-25, 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host, such as, yeast or *E. coli*).

In a preferred embodiment, a polynucleotide encoding an albumin portion of an albumin fusion protein of the invention is optimized for expression in yeast or mammalian cells. In further preferred embodiment, a polynucleotide encoding a Therapeutic protein portion of an albumin fusion protein of the invention is optimized for expression in yeast or mammalian cells. In a still further preferred embodiment, a polynucleotide encoding an albumin fusion protein of the invention is optimized for expression in yeast or mammalian cells.

In an alternative embodiment, a codon optimized polynucleotide encoding a Therapeutic protein portion of an albumin fusion protein of the invention does not hybridize to the wild type polynucleotide encoding the Therapeutic protein under stringent hybridization conditions as described herein. In a further embodiment, a codon optimized polynucleotide encoding an albumin portion of an albumin fusion protein of the invention does not hybridize to the wild type polynucleotide encoding the albumin protein under stringent hybridization conditions as described herein. In another embodiment, a codon optimized polynucleotide encoding an albumin fusion protein of the invention does not hybridize to the wild type polynucleotide encoding the Therapeutic protein portin or the albumin protein portion under stringent hybridization conditions as described herein.

In an additional embodiment, polynucleotides encoding a Therapeutic protein portion of an albumin fusion protein of the invention do not comprise, or alternatively consist of, the

naturally occurring sequence of that Therapeutic protein. In a further embodiment, polynucleotides encoding an albumin protein portion of an albumin fusion protein of the invention do not comprise, or alternatively consist of, the naturally occurring sequence of albumin protein. In an alternative embodiment, polynucleotides encoding an albumin fusion protein of the invention do not comprise, or alternatively consist of, the naturally occurring sequence of a Therapeutic protein portion or the albumin protein portion.

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985)). These allelic variants can vary at either the polynucleotide and/or polypeptide level and are included in the present invention. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the polypeptide of the present invention without substantial loss of biological function. As an example, Ron et al. (J. Biol. Chem. 268: 2984-2988 (1993)) reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem. 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods

described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which have a functional activity (e.g., biological activity and/or therapeutic activity). In highly preferred embodiments the invention provides variants of albumin fusion proteins that have a functional activity (e.g., biological activity and/or therapeutic activity, such as that disclosed in the "Biological Activity" column in Table 1) that corresponds to one or more biological and/or therapeutic activities of the Therapeutic protein corresponding to the Therapeutic protein portion of the albumin fusion protein. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity.

In preferred embodiments, the variants of the invention have conservative substitutions. By "conservative substitutions" is intended swaps within groups such as replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Guidance concerning how to make phenotypically silent amino acid substitutions is provided, for example, in Bowie et al., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. See Cunningham and Wells, Science 244:1081-1085 (1989). The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes

are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly. Besides conservative amino acid substitution, variants of the present invention include (i) polypeptides containing substitutions of one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) polypeptides containing substitutions of one or more of the amino acid residues having a substituent group, or (iii) polypeptides which have been fused with or chemically conjugated to another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), (iv) polypeptide containing additional amino acids, such as, for example, an IgG Fc fusion region peptide, . Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. See Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).

In specific embodiments, the polypeptides of the invention comprise, or alternatively, consist of, fragments or variants of the amino acid sequence of a Therapeutic protein described herein and/or human serum albumin, and/or albumin fusion protein of the invention, wherein the fragments or variants have 1-5, 5-10, 5-25, 5-50, 10-50 or 50-150, amino acid residue additions, substitutions, and/or deletions when compared to the reference amino acid sequence. In preferred embodiments, the amino acid substitutions are conservative. Nucleic acids encoding these polypeptides are also encompassed by the invention.

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as post-translational processing, or by chemical modification techniques which are well known in the art. Such modifications are well

described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POST-TRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth. Enzymol. 182:626-646 (1990); Rattan et al., Ann. N.Y. Acad. Sci. 663:48-62 (1992)).

Functional activity

"A polypeptide having functional activity" refers to a polypeptide capable of displaying one or more known functional activities associated with the full-length, pro-protein, and/or mature form of a Therapeutic protein. Such functional activities include, but are not limited to, biological activity, antigenicity [ability to bind (or compete with a polypeptide for binding) to an anti-polypeptide antibody], immunogenicity (ability to generate antibody which binds to a specific polypeptide of the invention), ability to form multimers with polypeptides of the invention, and ability to bind to a receptor or ligand for a polypeptide.

"A polypeptide having biological activity" refers to a polypeptide exhibiting activity similar to, but not necessarily identical to, an activity of a Therapeutic protein of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a

given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention).

In preferred embodiments, an albumin fusion protein of the invention has at least one biological and/or therapeutic activity associated with the Therapeutic protein (or fragment or variant thereof) when it is not fused to albumin.

The albumin fusion proteins of the invention can be assayed for functional activity (e.g., biological activity) using or routinely modifying assays known in the art, as well as assays described herein. Specifically, albumin fusion proteins may be assayed for functional activity (e.g., biological activity or therapeutic activity) using the assay referenced in the "Exemplary Activity Assay" column of Table 1. Additionally, one of skill in the art may routinely assay fragments of a Therapeutic protein corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention, for activity using assays referenced in its corresponding row of Table 1. Further, one of skill in the art may routinely assay fragments of an albumin protein corresponding to an albumin protein portion of an albumin fusion protein of the invention, for activity using assays known in the art and/or as described in the Examples section below.

For example, in one embodiment where one is assaying for the ability of an albumin fusion protein of the invention to bind or compete with a Therapeutic protein for binding to an anti-Therapeutic polypeptide antibody and/or anti-albumin antibody, various immunoassays known in the art can be used, including but not limited to, competitive and non-competitive assay systems using techniques such as radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitation reactions, immunodiffusion assays, in situ immunoassays (using colloidal gold, enzyme or radioisotope labels, for example), western blots, precipitation reactions, agglutination assays (e.g., gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. In one embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is labeled. Many means are known in the art for detecting binding in an immunoassay and are within the scope of the present invention.

In a preferred embodiment, where a binding partner (e.g., a receptor or a ligand) of a Therapeutic protein is identified, binding to that binding partner by an albumin fusion protein containing that Therapeutic protein as the Therapeutic protein portion of the fusion can be assayed, e.g., by means well-known in the art, such as, for example, reducing and non-

reducing gel chromatography, protein affinity chromatography, and affinity blotting. See generally, Phizicky et al., Microbiol. Rev. 59:94-123 (1995). In another embodiment, the ability of physiological correlates of an albumin fusion protein of the present invention to bind to a substrate(s) of the Therapeutic polypeptide corresponding to the Therapeutic portion of the albumin fusion protein of the invention can be routinely assayed using techniques known in the art.

In an alternative embodiment, where the ability of an albumin fusion protein of the invention to multimerize is being evaluated, association with other components of the multimer can be assayed, e.g., by means well-known in the art, such as, for example, reducing and non-reducing gel chromatography, protein affinity chromatography, and affinity blotting. See generally, Phizicky et al., *supra.*

In addition, assays described herein (see Examples and Table 1) and otherwise known in the art may routinely be applied to measure the ability of albumin fusion proteins of the present invention and fragments, variants and derivatives thereof to elicit biological activity and/or Therapeutic activity (either *in vitro* or *in vivo*) related to either the Therapeutic protein portion and/or albumin portion of the albumin fusion protein of the present invention. Other methods will be known to the skilled artisan and are within the scope of the invention.

Albumin

As described above, an albumin fusion protein of the invention comprises at least a fragment or variant of a Therapeutic protein and at least a fragment or variant of human serum albumin, which are associated with one another, preferably by genetic fusion or chemical conjugation.

The terms, human serum albumin (HSA) and human albumin (HA) are used interchangeably herein. The terms, "albumin and "serum albumin" are broader, and encompass human serum albumin (and fragments and variants thereof) as well as albumin from other species (and fragments and variants thereof).

As used herein, "albumin" refers collectively to albumin protein or amino acid sequence, or an albumin fragment or variant, having one or more functional activities (e.g., biological activities) of albumin. In particular, "albumin" refers to human albumin or fragments thereof (see EP 201 239, EP 322 094 WO 97/24445, WO95/23857) especially the mature form of human albumin as shown in Figure 15 and SEQ ID NO:18, or albumin from other vertebrates or fragments thereof, or analogs or variants of these molecules or fragments thereof.

In preferred embodiments, the human serum albumin protein used in the albumin fusion proteins of the invention contains one or both of the following sets of point mutations with reference to SEQ ID NO:18: Leu-407 to Ala, Leu-408 to Val, Val-409 to Ala; and Arg-

410 to Ala; or Arg-410 to A, Lys-413 to Gln, and Lys-414 to Gln (see, e.g., International Publication No. WO95/23857, hereby incorporated in its entirety by reference herein). In even more preferred embodiments, albumin fusion proteins of the invention that contain one or both of above-described sets of point mutations have improved stability/resistance to yeast Yap3p proteolytic cleavage, allowing increased production of recombinant albumin fusion proteins expressed in yeast host cells.

As used herein, a portion of albumin sufficient to prolong the therapeutic activity or shelf-life of the Therapeutic protein refers to a portion of albumin sufficient in length or structure to stabilize or prolong the therapeutic activity of the protein so that the shelf life of the Therapeutic protein portion of the albumin fusion protein is prolonged or extended compared to the shelf-life in the non-fusion state. The albumin portion of the albumin fusion proteins may comprise the full length of the HA sequence as described above or as shown in Figure 15, or may include one or more fragments thereof that are capable of stabilizing or prolonging the therapeutic activity. Such fragments may be of 10 or more amino acids in length or may include about 15, 20, 25, 30, 50, or more contiguous amino acids from the HA sequence or may include part or all of specific domains of HA. For instance, one or more fragments of HA spanning the first two immunoglobulin-like domains may be used.

The albumin portion of the albumin fusion proteins of the invention may be a variant of normal HA. The Therapeutic protein portion of the albumin fusion proteins of the invention may also be variants of the Therapeutic proteins as described herein. The term "variants" includes insertions, deletions and substitutions, either conservative or non conservative, where such changes do not substantially alter one or more of the oncotic, useful ligand-binding and non-immunogenic properties of albumin, or the active site, or active domain which confers the therapeutic activities of the Therapeutic proteins.

In particular, the albumin fusion proteins of the invention may include naturally occurring polymorphic variants of human albumin and fragments of human albumin, for example those fragments disclosed in EP 322 094 (namely HA (Pn), where n is 369 to 419). The albumin may be derived from any vertebrate, especially any mammal, for example human, cow, sheep, or pig. Non-mammalian albumins include, but are not limited to, hen and salmon. The albumin portion of the albumin fusion protein may be from a different animal than the Therapeutic protein portion.

Generally speaking, an HA fragment or variant will be at least 100 amino acids long, preferably at least 150 amino acids long. The HA variant may consist of or alternatively comprise at least one whole domain of HA, for example domains 1 (amino acids 1-194 of SEQ ID NO:18), 2 (amino acids 195-387 of SEQ ID NO:18), 3 (amino acids 388-585 of SEQ ID NO:18), 1 + 2 (1-387 of SEQ ID NO:18), 2 + 3 (195-585 of SEQ ID NO:18) or 1 + 3 (amino acids 1-194 of SEQ ID NO:18 + amino acids 388-585 of SEQ ID NO:18). Each

domain is itself made up of two homologous subdomains namely 1-105, 120-194, 195-291, 316-387, 388-491 and 512-585, with flexible inter-subdomain linker regions comprising residues Lys106 to Glu119, Glu292 to Val315 and Glu492 to Ala511.

Preferably, the albumin portion of an albumin fusion protein of the invention comprises at least one subdomain or domain of HA or conservative modifications thereof. If the fusion is based on subdomains, some or all of the adjacent linker is preferably used to link to the Therapeutic protein moiety.

Albumin Fusion Proteins

The present invention relates generally to albumin fusion proteins and methods of treating, preventing, or ameliorating diseases or disorders. As used herein, "albumin fusion protein" refers to a protein formed by the fusion of at least one molecule of albumin (or a fragment or variant thereof) to at least one molecule of a Therapeutic protein (or fragment or variant thereof). An albumin fusion protein of the invention comprises at least a fragment or variant of a Therapeutic protein and at least a fragment or variant of human serum albumin, which are associated with one another, preferably by genetic fusion (i.e., the albumin fusion protein is generated by translation of a nucleic acid in which a polynucleotide encoding all or a portion of a Therapeutic protein is joined in-frame with a polynucleotide encoding all or a portion of albumin) or chemical conjugation to one another. The Therapeutic protein and albumin protein, once part of the albumin fusion protein, may be referred to as a "portion", "region" or "moiety" of the albumin fusion protein.

In one embodiment, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a Therapeutic protein (e.g., as described in Table 1) and a serum albumin protein. In other embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a biologically active and/or therapeutically active fragment of a Therapeutic protein and a serum albumin protein. In other embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a biologically active and/or therapeutically active variant of a Therapeutic protein and a serum albumin protein. In preferred embodiments, the serum albumin protein component of the albumin fusion protein is the mature portion of serum albumin.

In further embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a Therapeutic protein, and a biologically active and/or therapeutically active fragment of serum albumin. In further embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a Therapeutic protein and a biologically active and/or therapeutically active variant of serum albumin. In preferred embodiments, the Therapeutic protein portion of the albumin fusion protein is the mature portion of the Therapeutic protein.

In further embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a biologically active and/or therapeutically active fragment or variant of a Therapeutic protein and a biologically active and/or therapeutically active fragment or variant of serum albumin. In preferred embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, the mature portion of a Therapeutic protein and the mature portion of serum albumin.

Preferably, the albumin fusion protein comprises HA as the N-terminal portion, and a Therapeutic protein as the C-terminal portion. Alternatively, an albumin fusion protein comprising HA as the C-terminal portion, and a Therapeutic protein as the N-terminal portion may also be used.

In other embodiments, the albumin fusion protein has a Therapeutic protein fused to both the N-terminus and the C-terminus of albumin. In a preferred embodiment, the Therapeutic proteins fused at the N- and C- termini are the same Therapeutic proteins. In a preferred embodiment, the Therapeutic proteins fused at the N- and C- termini are different Therapeutic proteins. In another preferred embodiment, the Therapeutic proteins fused at the N- and C- termini are different Therapeutic proteins which may be used to treat or prevent the same disease, disorder, or condition (e.g. as listed in the "Preferred Indication Y" column of Table 1). In another preferred embodiment, the Therapeutic proteins fused at the N- and C-termini are different Therapeutic proteins which may be used to treat or prevent diseases or disorders (e.g. as listed in the "Preferred Indication Y" column of Table 1) which are known in the art to commonly occur in patients simultaneously.

In addition to albumin fusion protein in which the albumin portion is fused N-terminal and/or C-terminal of the Therapeutic protein portion, albumin fusion proteins of the invention may also be produced by inserting the Therapeutic protein or peptide of interest (e.g., Therapeutic protein X as diclosed in Table 1) into an internal region of HA. For instance, within the protein sequence of the HA molecule a number of loops or turns exist between the end and beginning of α-helices, which are stabilized by disulphide bonds (see Figures 9-11). The loops, as determined from the crystal structure of HA (Fig. 13) (PDB identifiers 1AO6, 1BJ5, 1BKE, 1BM0, 1E7E to 1E7I and 1UOR) for the most part extend away from the body of the molecule. These loops are useful for the insertion, or internal fusion, of therapeutically active peptides, particularly those requiring a secondary structure to be functional, or Therapeutic proteins, to essentially generate an albumin molecule with specific biological activity.

Loops in human albumin structure into which peptides or polypeptides may be inserted to generate albumin fusion proteins of the invention include: Val54-Asn61, Thr76-Asp89, Ala92-Glu100, Gln170-Ala176, His247-Glu252, Glu266-Glu277, Glu280-His288, Ala362-Glu368, Lys439-Pro447,Val462-Lys475, Thr478-Pro486, and Lys560-Thr566. In

more preferred embodiments, peptides or polypeptides are inserted into the Val54-Asn61, Gln170-Ala176, and/or Lys560-Thr566 loops of mature human albumin (SEQ ID NO:18).

Peptides to be inserted may be derived from either phage display or synthetic peptide libraries screened for specific biological activity or from the active portions of a molecule with the desired function. Additionally, random peptide libraries may be generated within particular loops or by insertions of randomized peptides into particular loops of the HA molecule and in which all possible combinations of amino acids are represented.

Such library(s) could be generated on HA or domain fragments of HA by one of the following methods:

(a) randomized mutation of amino acids within one or more peptide loops of HA or HA domain fragments. Either one, more or all the residues within a loop could be mutated in this manner (for example see Fig. 10a);

(b) replacement of, or insertion into one or more loops of HA or HA domain fragments (*i.e.*, internal fusion) of a randomized peptide(s) of length X_n (where X is an amino acid and n is the number of residues (for example see Fig. 10b);

(c) N-, C- or N- and C- terminal peptide/protein fusions in addition to (a) and/or (b).

The HA or HA domain fragment may also be made multifunctional by grafting the peptides derived from different screens of different loops against different targets into the same HA or HA domain fragment.

In preferred embodiments, peptides inserted into a loop of human serum albumin are peptide fragments or peptide variants of the Therapeutic proteins disclosed in Table 1.. More particulary, the invention encompasses albumin fusion proteins which comprise peptide fragments or peptide variants at least 7 at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30, at least 35, or at least 40 amino acids in length inserted into a loop of human serum albumin. The invention also encompasses albumin fusion proteins which comprise peptide fragments or peptide variants at least 7 at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30, at least 35, or at least 40 amino acids fused to the N-terminus of human serum albumin. The invention also encompasses albumin fusion proteins which comprise peptide fragments or peptide variants at least 7 at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30, at least 35, or at least 40 amino acids fused to the C-terminus of human serum albumin.

Generally, the albumin fusion proteins of the invention may have one HA-derived region and one Therapeutic protein-derived region. Multiple regions of each protein, however, may be used to make an albumin fusion protein of the invention. Similarly, more than one Therapeutic protein may be used to make an albumin fusion protein of the invention.

For instance, a Therapeutic protein may be fused to both the N- and C-terminal ends of the HA. In such a configuration, the Therapeutic protein portions may be the same or different Therapeutic protein molecules. The structure of bifunctional albumin fusion proteins may be represented as: X-HA-Y or Y-HA-X.

For example, an anti-BLyS™ scFv-HA-IFNα-2b fusion may be prepared to modulate the immune response to IFNα-2b by anti-BLyS™ scFv. An alternative is making a bi (or even multi) functional dose of HA-fusions *e.g.* HA-IFNα-2b fusion mixed with HA-anti-BLyS™ scFv fusion or other HA-fusions in various ratio's depending on function, half-life etc.

Bi- or multi-functional albumin fusion proteins may also be prepared to target the Therapeutic protein portion of a fusion to a target organ or cell type via protein or peptide at the opposite terminus of HA.

As an alternative to the fusion of known therapeutic molecules, the peptides could be obtained by screening libraries constructed as fusions to the N-, C- or N- and C- termini of HA, or domain fragment of HA, of typically 6, 8, 12, 20 or 25 or X_n (where X is an amino acid (aa) and n equals the number of residues) randomized amino acids, and in which all possible combinations of amino acids were represented. A particular advantage of this approach is that the peptides may be selected *in situ* on the HA molecule and the properties of the peptide would therefore be as selected for rather than, potentially, modified as might be the case for a peptide derived by any other method then being attached to HA.

Additionally, the albumin fusion proteins of the invention may include a linker peptide between the fused portions to provide greater physical separation between the moieties and thus maximize the accessibility of the Therapeutic protein portion, for instance, for binding to its cognate receptor. The linker peptide may consist of amino acids such that it is flexible or more rigid.

The linker sequence may be cleavable by a protease or chemically to yield the growth hormone related moiety. Preferably, the protease is one which is produced naturally by the host, for example the *S. cerevisiae* protease *kex2* or equivalent proteases.

Therefore, as described above, the albumin fusion proteins of the invention may have the following formula R1-L-R2; R2-L-R1; or R1-L-R2-L-R1, wherein R1 is at least one Therapeutic protein, peptide or polypeptide sequence, and not necessarily the same Therapeutic protein, L is a linker and R2 is a serum albumin sequence.

In preferred embodiments, Albumin fusion proteins of the invention comprising a Therapeutic protein have extended shelf life compared to the shelf life the same Therapeutic protein when not fused to albumin. Shelf-life typically refers to the time period over which the therapeutic activity of a Therapeutic protein in solution or in some other storage

formulation, is stable without undue loss of therapeutic activity. Many of the Therapeutic proteins are highly labile in their unfused state. As described below, the typical shelf-life of these Therapeutic proteins is markedly prolonged upon incorporation into the albumin fusion protein of the invention.

Albumin fusion proteins of the invention with "prolonged" or "extended" shelf-life exhibit greater therapeutic activity relative to a standard that has been subjected to the same storage and handling conditions. The standard may be the unfused full-length Therapeutic protein. When the Therapeutic protein portion of the albumin fusion protein is an analog, a variant, or is otherwise altered or does not include the complete sequence for that protein, the prolongation of therapeutic activity may alternatively be compared to the unfused equivalent of that analog, variant, altered peptide or incomplete sequence. As an example, an albumin fusion protein of the invention may retain greater than about 100% of the therapeutic activity, or greater than about 105%, 110%, 120%, 130%, 150% or 200% of the therapeutic activity of a standard when subjected to the same storage and handling conditions as the standard when compared at a given time point.

Shelf-life may also be assessed in terms of therapeutic activity remaining after storage, normalized to therapeutic activity when storage began. Albumin fusion proteins of the invention with prolonged or extended shelf-life as exhibited by prolonged or extended therapeutic activity may retain greater than about 50% of the therapeutic activity, about 60%, 70%, 80%, or 90% or more of the therapeutic activity of the equivalent unfused Therapeutic protein when subjected to the same conditions. For example, as discussed in Example 1, an albumin fusion protein of the invention comprising hGH fused to the full length HA sequence may retain about 80% or more of its original activity in solution for periods of up to 5 weeks or more under various temperature conditions.

Expression of Fusion Proteins

The albumin fusion proteins of the invention may be produced as recombinant molecules by secretion from yeast, a microorganism such as a bacterium, or a human or animal cell line. Preferably, the polypeptide is secreted from the host cells. We have found that, by fusing the hGH coding sequence to the HA coding sequence, either to the 5' end or 3' end, it is possible to secrete the albumin fusion protein from yeast without the requirement for a yeast-derived pro sequence. This was surprising, as other workers have found that a yeast derived pro sequence was needed for efficient secretion of hGH in yeast.

For example, Hiramatsu *et al.* (Appl Environ Microbiol 56:2125 (1990); Appl Environ Microbiol 57:2052 (1991)) found that the N-terminal portion of the pro sequence in the *Mucor pusillus* rennin pre-pro leader was important. Other authors, using the MFα-1 signal, have always included the MFα-1 pro sequence when secreting hGH. The pro sequences were

believed to assist in the folding of the hGH by acting as an intramolecular chaperone. The present invention shows that HA or fragments of HA can perform a similar function.

Hence, a particular embodiment of the invention comprises a DNA construct encoding a signal sequence effective for directing secretion in yeast, particularly a yeast-derived signal sequence (especially one which is homologous to the yeast host), and the fused molecule of the first aspect of the invention, there being no yeast-derived pro sequence between the signal and the mature polypeptide.

The *Saccharomyces cerevisiae* invertase signal is a preferred example of a yeast-derived signal sequence.

Conjugates of the kind prepared by Poznansky *et al.*, (FEBS Lett. 239:18 (1988)), in which separately-prepared polypeptides are joined by chemical cross-linking, are not contemplated.

The present invention also includes a cell, preferably a yeast cell transformed to express an albumin fusion protein of the invention. In addition to the transformed host cells themselves, the present invention also contemplates a culture of those cells, preferably a monoclonal (clonally homogeneous) culture, or a culture derived from a monoclonal culture, in a nutrient medium. If the polypeptide is secreted, the medium will contain the polypeptide, with the cells, or without the cells if they have been filtered or centrifuged away. Many expression systems are known and may be used, including bacteria (for example *E. coli* and *Bacillus subtilis*), yeasts (for example *Saccharomyces cerevisiae*, *Kluyveromyces lactis* and *Pichia pastoris*, filamentous fungi (for example *Aspergillus*), plant cells, animal cells and insect cells.

Preferred yeast strains to be used in the production of albumin fusion proteins are D88, DXY1 and BXP10. D88 [*leu2-3*, *leu2-122*, *can1*, *pra1*, *ubc4*] is a derivative of parent strain AH22*his*$^+$ (also known as DB1; see, e.g., Sleep *et al.* Biotechnology 8:42-46 (1990)). The strain contains a *leu2* mutation which allows for auxotropic selection of 2 micron-based plasmids that contain the LEU2 gene. D88 also exhibits a derepression of PRB1 in glucose excess. The PRB1 promoter is normally controlled by two checkpoints that monitor glucose levels and growth stage. The promoter is activated in wild type yeast upon glucose depletion and entry into stationary phase. Strain D88 exhibits the repression by glucose but maintains the induction upon entry into stationary phase. The PRA1 gene encodes a yeast vacuolar protease, YscA endoprotease A, that is localized in the ER. The UBC4 gene is in the ubiquitination pathway and is involved in targeting short lived and abnormal proteins for ubiquitin dependant degradation. Isolation of this ubc4 mutation was found to increase the copy number of an expression plasmid in the cell and cause an increased level of expression of a desired protein expressed from the plasmid (see, e.g., International Publication No. WO99/00504, hereby incorporated in its entirety by reference herein).

DXY1, a derivative of D88, has the following genotype: [*leu2-3*, *leu2-122*, *can1*, *pra1*, *ubc4*, *ura3::yap3*]. In addition to the mutations isolated in D88, this strain also has a knockout of the YAP3 protease. This protease causes cleavage of mostly di-basic residues (RR, RK, KR, KK) but can also promote cleavage at single basic residues in proteins. Isolation of this yap3 mutation resulted in higher levels of full length HSA production (see, e.g., U.S. Patent No. 5,965,386, and Kerry-Williams et al., Yeast 14:161-169 (1998), hereby incorporated in their entireties by reference herein).

BXP10 has the following genotype: *leu2-3*, *leu2-122*, *can1*, *pra1*, *ubc4*, *ura3*, *yap3::URA3*, *lys2*, *hsp150::LYS2*, *pmt1::URA3*. In addition to the mutations isolated in DXY1, this strain also has a knockout of the PMT1 gene and the HSP150 gene. The PMT1 gene is a member of the evolutionarily conserved family of dolichyl-phosphate-D-mannose protein O-mannosyltransferases (Pmts). The transmembrane topology of Pmt1p suggests that it is an integral membrane protein of the endoplasmic reticulum with a role in O-linked glycosylation. This mutation serves to reduce/eliminate O-linked glycosylation of HSA fusions (see, e.g., International Publication No. WO00/44772, hereby incorporated in its entirety by reference herein). Studies revealed that the Hsp150 protein is inefficiently separated from rHA by ion exchange chromatography. The mutation in the HSP150 gene removes a potential contaminant that has proven difficult to remove by standard purification techniques. See, e.g., U.S. Patent No. 5,783,423, hereby incorporated in its entirety by reference herein.

The desired protein is produced in conventional ways, for example from a coding sequence inserted in the host chromosome or on a free plasmid. The yeasts are transformed with a coding sequence for the desired protein in any of the usual ways, for example electroporation. Methods for transformation of yeast by electroporation are disclosed in Becker & Guarente (1990) *Methods Enzymol.* 194, 182.

Successfully transformed cells, *i.e.*, cells that contain a DNA construct of the present invention, can be identified by well known techniques. For example, cells resulting from the introduction of an expression construct can be grown to produce the desired polypeptide. Cells can be harvested and lysed and their DNA content examined for the presence of the DNA using a method such as that described by Southern (1975) *J. Mol. Biol.* 98, 503 or Berent *et al.* (1985) *Biotech.* 3, 208. Alternatively, the presence of the protein in the supernatant can be detected using antibodies.

Useful yeast plasmid vectors include pRS403-406 and pRS413-416 and are generally available from Stratagene Cloning Systems, La Jolla, CA 92037, USA. Plasmids pRS403, pRS404, pRS405 and pRS406 are Yeast Integrating plasmids (YIps) and incorporate the yeast selectable markers HIS3, 7RP1, LEU2 and URA3. Plasmids pRS413-416 are Yeast Centromere plasmids (Ycps).

Preferred vectors for making albumin fusion proteins for expression in yeast include pPPC0005, pScCHSA, pScNHSA, and pC4:HSA which are described in detail in Example 2. Figure 4 shows a map of the pPPC0005 plasmid that can be used as the base vector into which polynucleotides encoding Therapeutic proteins may be cloned to form HA-fusions. It contains a *PRB1 S. cerevisiae* promoter (PRB1p), a Fusion leader sequence (FL), DNA encoding HA (rHA) and an *ADH1 S. cerevisiae* terminator sequence. The sequence of the fusion leader sequence consists of the first 19 amino acids of the signal peptide of human serum albumin (SEQ ID NO:29) and the last five amino acids of the mating factor alpha 1 promoter (SLDKR, see EP-A-387 319 which is hereby incorporated by reference in its entirety.

The plasmids, pPPC0005, pScCHSA, pScNHSA, and pC4:HSA were deposited on April 11, 2001 at the American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia 20110-2209 and given accession numbers ATCC ______, ______, ______, and ______, respectively. Another vector useful for expressing an albumin fusion protein in yeast the pSAC35 vector which is described in Sleep *et al.*, BioTechnology 8:42 (1990) which is hereby incorporated by reference in its entirety.

A variety of methods have been developed to operably link DNA to vectors via complementary cohesive termini. For instance, complementary homopolymer tracts can be added to the DNA segment to be inserted to the vector DNA. The vector and DNA segment are then joined by hydrogen bonding between the complementary homopolymeric tails to form recombinant DNA molecules.

Synthetic linkers containing one or more restriction sites provide an alternative method of joining the DNA segment to vectors. The DNA segment, generated by endonuclease restriction digestion, is treated with bacteriophage T4 DNA polymerase or E. coli DNA polymerase I, enzymes that remove protruding, γ-single-stranded termini with their 3' 5'-exonucleolytic activities, and fill in recessed 3'-ends with their polymerizing activities.

The combination of these activities therefore generates blunt-ended DNA segments. The blunt-ended segments are then incubated with a large molar excess of linker molecules in the presence of an enzyme that is able to catalyze the ligation of blunt-ended DNA molecules, such as bacteriophage T4 DNA ligase. Thus, the products of the reaction are DNA segments carrying polymeric linker sequences at their ends. These DNA segments are then cleaved with the appropriate restriction enzyme and ligated to an expression vector that has been cleaved with an enzyme that produces termini compatible with those of the DNA segment.

Synthetic linkers containing a variety of restriction endonuclease sites are commercially available from a number of sources including International Biotechnologies Inc, New Haven, CT, USA.

A desirable way to modify the DNA in accordance with the invention, if, for example,

HA variants are to be prepared, is to use the polymerase chain reaction as disclosed by Saiki *et al.* (1988) *Science* 239, 487-491. In this method the DNA to be enzymatically amplified is flanked by two specific oligonucleotide primers which themselves become incorporated into the amplified DNA. The specific primers may contain restriction endonuclease recognition sites which can be used for cloning into expression vectors using methods known in the art.

Exemplary genera of yeast contemplated to be useful in the practice of the present invention as hosts for expressing the albumin fusion proteins are *Pichia* (formerly classified as Hansenula), *Saccharomyces, Kluyveromyces, Aspergillus, Candida, Torulopsis, Torulaspora, Schizosaccharomyces, Citeromyces, Pachysolen, Zygosaccharomyces, Debaromyces, Trichoderma, Cephalosporium, Humicola, Mucor, Neurospora, Yarrowia, Metschunikowia, Rhodosporidium, Leucosporidium, Botryoascus, Sporidiobolus, Endomycopsis*, and the like. Preferred genera are those selected from the group consisting of *Saccharomyces, Schizosaccharomyces, Kluyveromyces, Pichia* and *Torulaspora*. Examples of *Saccharomyces spp*. are *S. cerevisiae, S. italicus and S. rouxii.*

Examples of *Kluyveromyces* spp. are *K. fragilis, K. lactis and K. marxianus*. A suitable *Torulaspora* species is *T. delbrueckii*. Examples of *Pichia* (Hansenula) spp. are *P. angusta* (formerly *H. polymorpha*), *P. anomala* (formerly *H. anomala*) and *P. pastoris*. Methods for the transformation of *S. cerevisiae* are taught generally in EP 251 744, EP 258 067 and WO 90/01063, all of which are incorporated herein by reference.

Preferred exemplary species of *Saccharomyces* include *S. cerevisiae, S. italicus, S. diastaticus*, and *Zygosaccharomyces rouxii*. Preferred exemplary species of *Kluyveromyces* include *K. fragilis* and *K. lactis*. Preferred exemplary species of *Hansenula* include *H. polymorpha* (now *Pichia angusta*), *H. anomala* (now *Pichia anomala*), and *Pichia capsulata*. Additional preferred exemplary species of *Pichia* include *P. pastoris*. Preferred exemplary species of *Aspergillus* include *A. niger* and *A. nidulans*. Preferred exemplary species of *Yarrowia* include *Y. lipolytica*. Many preferred yeast species are available from the ATCC. For example, the following preferred yeast species are available from the ATCC and are useful in the expression of albumin fusion proteins: *Saccharomyces cerevisiae* Hansen, teleomorph strain BY4743 *yap3* mutant (ATCC Accession No. 4022731); *Saccharomyces cerevisiae* Hansen, teleomorph strain BY4743 *hsp150* mutant (ATCC Accession No. 4021266); *Saccharomyces cerevisiae* Hansen, teleomorph strain BY4743 *pmt1* mutant (ATCC Accession No. 4023792); *Saccharomyces cerevisiae* Hansen, teleomorph (ATCC Accession Nos. 20626; 44773; 44774; and 62995); *Saccharomyces diastaticus* Andrews et Gilliland ex van der Walt, teleomorph (ATCC Accession No. 62987); *Kluyveromyces lactis* (Dombrowski) van der Walt, teleomorph (ATCC Accession No. 76492); *Pichia angusta* (Teunisson et al.) Kurtzman, teleomorph deposited as *Hansenula polymorpha* de Morais et Maia, teleomorph (ATCC Accession No. 26012); *Aspergillus niger* van Tieghem, anamorph

(ATCC Accession No. 9029); *Aspergillus niger* van Tieghem, anamorph (ATCC Accession No. 16404); *Aspergillus nidulans* (Eidam) Winter, anamorph (ATCC Accession No. 48756); and *Yarrowia lipolytica* (Wickerham et al.) van der Walt et von Arx, teleomorph (ATCC Accession No. 201847).

Suitable promoters for S. cerevisiae include those associated with the PGKI gene, GAL1 or GAL10 genes, CYCI, PHO5, TRPI, ADHI, ADH2, the genes for glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, triose phosphate isomerase, phosphoglucose isomerase, glucokinase, alpha-mating factor pheromone, [a mating factor pheromone], the PRBI promoter, the GUT2 promoter, the GPDI promoter, and hybrid promoters involving hybrids of parts of 5' regulatory regions with parts of 5' regulatory regions of other promoters or with upstream activation sites (*e.g.* the promoter of EP-A-258 067).

Convenient regulatable promoters for use in *Schizosaccharomyces pombe* are the thiamine-repressible promoter from the nmt gene as described by Maundrell (1990) *J. Biol. Chem.* 265, 10857-10864 and the glucose repressible jbpl gene promoter as described by Hoffman & Winston (1990) *Genetics* 124, 807-816.

Methods of transforming *Pichia* for expression of foreign genes are taught in, for example, Cregg *et al.* (1993), and various Phillips patents (*e.g.* US 4 857 467, incorporated herein by reference), and *Pichia* expression kits are commercially available from Invitrogen BV, Leek, Netherlands, and Invitrogen Corp., San Diego, California. Suitable promoters include AOXI and AOX2. Gleeson *et al.* (1986) J. Gen. Microbiol. 132, 3459-3465 include information on *Hansenula* vectors and transformation, suitable promoters being MOX1 and FMD1; whilst EP 361 991, Fleer *et al.* (1991) and other- publications from Rhone-Poulenc Rorer teach how to express foreign proteins in *Kluyveromyces* spp., a suitable promoter being PGKI.

The transcription termination signal is preferably the 3' flanking sequence of a eukaryotic gene which contains proper signals for transcription termination and polyadenylation. Suitable 3' flanking sequences may, for example, be those of the gene naturally linked to the expression control sequence used, *i.e.* may correspond to the promoter. Alternatively, they may be different in which case the termination signal of the S. cerevisiae ADHI gene is preferred.

The desired albumin fusion protein may be initially expressed with a secretion leader sequence, which may be any leader effective in the yeast chosen. Leaders useful in S. cerevisiae include that from the mating factor α polypeptide (MF α-1) and the hybrid leaders of EP-A-387 319. Such leaders (or signals) are cleaved by the yeast before the mature albumin is released into the surrounding medium. Further such leaders include those of *S. cerevisiae* invertase (SUC2) disclosed in JP 62-096086 (granted as 911036516), acid

phosphatase (PH05), the pre-sequence of MFα-1, 0 glucanase (BGL2) and killer toxin; *S. diastaticus* glucoarnylase II; *S. carlsbergensis* α-galactosidase (MEL1); *K. lactis* killer toxin; and *Candida glucoarnylase*.

Additional Methods of Recombinant and Synthetic Production of Albumin Fusion Proteins

The present invention also relates to vectors containing a polynucleotide encoding an albumin fusion protein of the present invention, host cells, and the production of albumin fusion proteins by synthetic and recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides encoding albumin fusion proteins of the invention may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli lac, trp, phoA* and *tac* promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418, glutamine synthase, or neomycin resistance for eukaryotic cell culture, and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli,* Streptomyces and *Salmonella typhimurium* cells; fungal cells, such as yeast cells (e.g., *Saccharomyces cerevisiae* or *Pichia pastoris* (ATCC Accession No. 201178)); insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS,NSO, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9,